

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 15:13:17 ; Search time 22 Seconds
(without alignments)
14.080 Million cell updates/sec

Title: US-09-848-834A-6
Perfect score: 29
Sequence: 1 SSGPSL 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 40703

Minimum DB seq length: 0
Maximum DB seq length: 6

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	72.4	6	1 US-08-447-500-25	Sequence 25, Appl
2	21	72.4	6	1 US-08-454-097-25	Sequence 25, Appl
3	21	72.4	6	1 US-08-453-866-25	Sequence 25, Appl
4	21	72.4	6	2 US-08-460-502-4	Sequence 4, Appl
5	21	72.4	6	3 US-08-185-359-25	Sequence 25, Appl
6	21	72.4	6	4 US-08-148-711A-4	Sequence 4, Appl
7	19	65.5	5	4 US-08-891-525-7	Sequence 7, Appl
8	18	62.1	4	1 US-08-432-617-6	Sequence 6, Appl
9	18	62.1	4	1 US-08-206-789-2	Sequence 2, Appl
10	18	62.1	4	1 US-08-358-160-173	Sequence 173, App
11	18	62.1	4	1 US-08-358-160-174	Sequence 5, Appl
12	18	62.1	5	1 US-08-432-617-5	Sequence 20, Appl
13	18	62.1	5	4 US-08-557-050-20	Sequence 4, Appl
14	18	62.1	6	1 US-08-432-617-4	Sequence 19, Appl
15	18	62.1	6	3 US-08-274-642-19	Sequence 22, Appl
16	18	62.1	6	4 US-08-557-050-22	Sequence 23, Appl
17	18	62.1	6	4 US-08-557-050-23	Sequence 23, Appl
18	17	58.6	4	2 US-08-248-839C-142	Sequence 142, App
19	17	58.6	4	3 US-08-435-568A-17	Sequence 17, Appl
20	17	58.6	4	3 US-08-330-970-18	Sequence 18, Appl
21	17	58.6	5	1 US-07-990-301A-9	Sequence 9, Appl
22	17	58.6	5	1 US-08-076-032-66	Sequence 66, Appl
23	17	58.6	5	1 US-08-225-224-51	Sequence 51, Appl
24	17	58.6	5	1 US-08-340-428B-46	Sequence 46, Appl
25	17	58.6	5	2 US-08-730-486-66	Sequence 66, Appl
26	17	58.6	5	2 US-08-755-728-6	Sequence 6, Appl
27	17	58.6	5	2 US-08-974-655-6	Sequence 6, Appl

28	17	58.6	5	3 US-08-722-258-51	Sequence 51, Appl
29	17	58.6	5	3 US-09-283-011-6	Sequence 6, Appl
30	17	58.6	5	3 US-08-931-858E-224	Sequence 224, App
31	17	58.6	5	4 US-09-440-344-3	Sequence 3, Appl
32	17	58.6	5	5 PCT-US93-07306-46	Sequence 46, Appl
33	17	58.6	5	5 PCT-US95-04468-51	Sequence 51, Appl
34	17	58.6	5	6 520320-26	Patent No. 5200320
35	17	58.6	5	6 5217869-34	Patent No. 5217869
36	17	58.6	6	1 US-07-990-301A-12	Sequence 12, Appl
37	17	58.6	6	1 US-08-225-224-57	Sequence 57, Appl
38	17	58.6	6	1 US-08-403-378B-20	Sequence 20, Appl
39	17	58.6	6	2 US-08-353-476-102	Sequence 102, App
40	17	58.6	6	2 US-08-760-903-9	Sequence 9, Appl
41	17	58.6	6	3 US-08-722-258-57	Sequence 57, Appl
42	17	58.6	6	4 US-08-557-050-18	Sequence 18, Appl
43	17	58.6	6	4 US-08-557-050-19	Sequence 19, Appl
44	17	58.6	6	4 US-08-557-050-21	Sequence 21, Appl
45	17	58.6	6	4 US-08-757-425B-43	Sequence 43, Appl

ALIGNMENTS

RESULT 1
US-08-447-500-25
; Sequence 25, Application US/08447500
; Patent No. 5627064
; GENERAL INFORMATION:
; APPLICANT: Hoeckstra, Merl F.
; TITLE OF INVENTION: PROTEIN KINASES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1980 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,500
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/008,001
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-2458
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; CLONE: Protein Kinase
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..6
; US-08-447-500-25

Query Match 72.4%; Score 21; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GPSL 6
Db 1 GPSL 4

RESULT 2

US-08-454-097-25
; Sequence 25, Application US/08454097
; Patent No. 5686412
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; TITLE OF INVENTION: Protein Kinases
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/454,097
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/185,359
; FILING DATE: 21-JAN-1994
; APPLICATION NUMBER: US 08/008,001
; FILING DATE: 21-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,783
; FILING DATE: 03-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5686412and, Greta B.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31853
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; CLONE: Protein Kinase
; NAME/KEY: Peptide
; LOCATION: 1..6
US-08-454-097-25

Query Match 72.4%; Score 21; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GPSL 6
Db 1 GPSL 4

RESULT 3

US-08-453-866-25
; Sequence 25, Application US/08453866

Patent No. 5756289
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; TITLE OF INVENTION: PROTEIN KINASES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/453,866
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,001
; FILING DATE: 20-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-2458
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; CLONE: Protein Kinase
; NAME/KEY: Peptide
; LOCATION: 1..6
US-08-453-866-25

Query Match 72.4%; Score 21; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GPSL 6
Db 1 GPSL 4

RESULT 4

US-08-460-502-4
; Sequence 4, Application US/08460502
; Patent No. 5843464
; GENERAL INFORMATION:
; APPLICANT: Bakaletz, Lauren O.
; APPLICANT: Kaumaya, Parvin T.
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calfee, Halter and Griswold
; STREET: 800 Superior Avenue
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 44114-2688
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,502
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Golrick, Mary E.
; REGISTRATION NUMBER: 34,829
; REFERENCE/DOCKET NUMBER: 22727/00120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (216) 622-8458
; TELEFAX: (216) 241-0816
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-460-502-4

Query Match 72.4%; Score 21; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPSL 6
DB 1 GPSL 4

RESULT 5
US-08-185-359-25
; Sequence 25, Application US/08185359
; Patent No. 6060296
; GENERAL INFORMATION:
; APPLICANT: Hoeckstra, Merl F.
; TITLE OF INVENTION: Protein Kinases
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60608-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,359
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,001
; FILING DATE: 21-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,783
; FILING DATE: 03-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6060296and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31853
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; CLONE: Protein Kinase
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..6
US-08-185-359-25

Query Match 72.4%; Score 21; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPSL 6
DB 1 GPSL 4

RESULT 6
US-09-148-711A-4
; Sequence 4, Application US/09148711A
; Patent No. 6436405
; GENERAL INFORMATION:
; APPLICANT: The Ohio State University
; TITLE OF INVENTION: Synthetic Chimeric Fimbrin Peptides
; FILE REFERENCE: 18525-04010
; CURRENT APPLICATION NUMBER: US/09/148,711A
; CURRENT FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 08/460,502
; PRIOR FILING DATE: 1995-06-02
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 4
; LENGTH: 6
; TYPE: PRT
; ORGANISM: synthetic construct
US-09-148-711A-4

Query Match 72.4%; Score 21; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPSL 6
DB 1 GPSL 4

RESULT 7
US-08-891-525-7
; Sequence 7, Application US/08891525
; Patent No. 6558900
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; APPLICANT: Liu, Xuesong
; TITLE OF INVENTION: Regulation of Apoptosis and In Vitro
; TITLE OF INVENTION: Model for Studies thereof
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,525
; FILING DATE: 11-JUL-1997
```

```
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/021,268
/ FILING DATE: 12-JUL-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Feiber, Donna M.
/ REGISTRATION NUMBER: 33,878
/ REFERENCE/DOCKET NUMBER: 45-96
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (303) 499-8080
/ TELEFAX: (303) 499-8089
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 5 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: not relevant
/ MOLECULE TYPE: peptide
/ HYPOTHETICAL: NO
/ FRAGMENT TYPE: internal
/ US-08-891-525-7

Query Match 65.5%; Score 19; DB 4; Length 5;
Best Local Similarity 60.0%; Pred. No. 3e+05; 0; Indels 0;
Matches 3; Conservative 2; Mismatches 0; Gaps 0;

QY 2 SGPSL 6
DB 1 TGNL 5

RESULT 8
US-08-432-617-6
/ Sequence 6, Application US/08432617
/ Patent No. 5495000
/ GENERAL INFORMATION:
/ APPLICANT: Krstenansky, John L
/ TITLE OF INVENTION: Anticoagulant Peptides
/ NUMBER OF SEQUENCES: 14
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Marion Merrell Dow Inc.
/ STREET: 2110 East Galbraith Rd.
/ CITY: Cincinnati P. O. Box 156300
/ STATE: Ohio
/ COUNTRY: USA
/ ZIP: 45215-6300
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/432,617
/ FILING DATE:
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/971,909
/ FILING DATE: 18-DEC-1992
/ APPLICATION NUMBER: US 07/557,288
/ FILING DATE: 24-JUL-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: WO PCT/US91/04558
/ FILING DATE: 28-JUN-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Boudreaux, William R
/ REGISTRATION NUMBER: 35,796
/ REFERENCE/DOCKET NUMBER: M01556 US-A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (513) 948-6566
/ TELEFAX: (513) 948-7961
/ TELEX: 214320
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FEATURE:
/ NAME/KEY: Peptide
/ LOCATION: 1..4
/ US-08-206-789-2

Query Match 62.1%; Score 18; DB 1; Length 4;
Best Local Similarity 75.0%; Pred. No. 3e+05; 0; Indels 0;
Matches 3; Conservative 1; Mismatches 0; Gaps 0;

QY 3 GPSL 6
DB 1 GPAL 4

RESULT 10
US-08-432-617-6
/ LENGTH: 4 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-08-432-617-6

Query Match 62.1%; Score 18; DB 1; Length 4;
Best Local Similarity 75.0%; Pred. No. 3e+05; 0; Indels 0;
Matches 3; Conservative 1; Mismatches 0; Gaps 0;

QY 1 SSGP 4
DB 1 TSQP 4

RESULT 9
US-08-206-789-2
/ Sequence 2, Application US/08206789
/ Patent No. 558854
/ GENERAL INFORMATION:
/ APPLICANT: Orlowski, Marian
/ APPLICANT: Cardozo, Christopher
/ APPLICANT: Vinitzky, Alexander
/ TITLE OF INVENTION: SUBSTRATE-RELATED
/ TITLE OF INVENTION: PEPTIDYL-ALDEHYDE
/ TITLE OF INVENTION: INHIBITORS OF THE PROTEOLYTIC ACTIVITY OF THE
/ TITLE OF INVENTION: MULTICATALYTIC PROTEINASE COMPLEX
/ NUMBER OF SEQUENCES: 5
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: BRUMBAUGH, GRAVES, DONOHUE &
/ ADDRESS: RAYMOND
/ STREET: 30 Rockefeller Plaza
/ CITY: New York
/ STATE: New York
/ COUNTRY: USA
/ ZIP: 10112-2500
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/206,789
/ FILING DATE:
/ CLASSIFICATION: 530
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Seide, Rochelle K.
/ REGISTRATION NUMBER: 32,300
/ REFERENCE/DOCKET NUMBER: A29525 - 165/25989
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-408-2626
/ TELEFAX: 212-765-2519
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FEATURE:
/ NAME/KEY: Peptide
/ LOCATION: 1..4
/ US-08-206-789-2
```


US-08-358-160-173
; Sequence 173, Application US/08358160
; Patent No. 5663143
; GENERAL INFORMATION:
; APPLICANT: LEY, Arthur C.
; APPLICANT: LADNER, Robert C.
; APPLICANT: GUTERMAN, Sonia K.
; APPLICANT: ROBERTS, Bruce L.
; APPLICANT: MARKLAND, William
; APPLICANT: KENT, Rachel B.
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; TITLE OF INVENTION: DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
; NUMBER OF SEQUENCES: 234
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W. Suite 300
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/358,160
; FILING DATE: 16-DEC-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/133,031
; FILING DATE: 13-OCT-1993
; APPLICATION NUMBER: US 08/009,319
; FILING DATE: 26-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/664,989
; FILING DATE: 01-MAR-1991
; APPLICATION NUMBER: US 07/487,063
; FILING DATE: 02-MAR-1990
; APPLICATION NUMBER: US 07/240,160
; FILING DATE: 02-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Cooper, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: LEY=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 173:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-358-160-173
Query Match 62.1%; Score 18; DB 1; Length 4;
Best Local Similarity 75.0%; Pred. No. 3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SSGP 4
Db 1 SAGP 4

RESULT 11
US-08-358-160-174
; Sequence 174, Application US/08358160

Patent No. 5663143
; GENERAL INFORMATION:
; APPLICANT: LEY, Arthur C.
; APPLICANT: LADNER, Robert C.
; APPLICANT: GUTERMAN, Sonia K.
; APPLICANT: ROBERTS, Bruce L.
; APPLICANT: MARKLAND, William
; APPLICANT: KENT, Rachel B.
; TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
; TITLE OF INVENTION: DOMAINS THAT INHIBIT HUMAN NEUTROPHIL ELASTASE
; NUMBER OF SEQUENCES: 234
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W. Suite 300
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/358,160
; FILING DATE: 16-DEC-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/133,031
; FILING DATE: 13-OCT-1993
; APPLICATION NUMBER: US 08/009,319
; FILING DATE: 26-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/664,989
; FILING DATE: 01-MAR-1991
; APPLICATION NUMBER: US 07/487,063
; FILING DATE: 02-MAR-1990
; APPLICATION NUMBER: US 07/240,160
; FILING DATE: 02-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Cooper, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: LEY=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 174:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-358-160-174
Query Match 62.1%; Score 18; DB 1; Length 4;
Best Local Similarity 75.0%; Pred. No. 3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SSGP 4
Db 1 STGP 4
RESULT 12
US-08-432-617-5
; Sequence 5, Application US/08432617
; Patent No. 5495000
; GENERAL INFORMATION:

```

; APPLICANT: Krstenansky, John L
; TITLE OF INVENTION: Anticoagulant Peptides
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marion Merrell Dow Inc.
; STREET: 2110 East Galbraith Rd.
; CITY: Cincinnati P. O. Box 156300
; STATE: Ohio
; COUNTRY: USA
; ZIP: 45215-6300
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,617
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/971,909
; FILING DATE: 18-DEC-1992
; APPLICATION NUMBER: US 07/557,288
; FILING DATE: 24-JUL-1990
; PRIOR APPLICATION DATA: WO PCT/US91/04658
; FILING DATE: 28-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Boudreaux, William R
; REGISTRATION NUMBER: 35,796
; REFERENCE/DOCKET NUMBER: M01556 US-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (513) 948-6566
; TELEFAX: (513) 948-7961
; TELEX: 214320
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-432-617-5

Query Match 62.1%; Score 18; DB 1; Length 5;
Best Local Similarity 75.0%; Pred. No. 3e-05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGP 4
Db 2 TSGP 5

RESULT 13
US-08-557-050-20
; Sequence 20, Application US/08557050
; Patent No. 6491916
; GENERAL INFORMATION:
; APPLICANT: Bluestone, Jeffrey A.
; APPLICANT: Zivin, Robert A.
; APPLICANT: Jolliffe, Linda K.
; TITLE OF INVENTION: METHODS AND MATERIALS FOR MODULATION OF
; TITLE OF INVENTION: THE IMMUNO-SUPPRESSIVE ACTIVITY AND TOXICITY OF MONOCLONAL
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,050
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/06198
; FILING DATE: 01-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/070,116
; FILING DATE: 01-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: ARCD:208
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-557-050-20

Query Match 62.1%; Score 18; DB 4; Length 5;
Best Local Similarity 75.0%; Pred. No. 3e-05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SGPS 5
Db 2 AGPS 5

RESULT 14
US-08-432-617-4
; Sequence 4, Application US/08432617
; Patent No. 5495000
; GENERAL INFORMATION:
; APPLICANT: Krstenansky, John L
; TITLE OF INVENTION: Anticoagulant Peptides
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marion Merrell Dow Inc.
; STREET: 2110 East Galbraith Rd.
; CITY: Cincinnati P. O. Box 156300
; STATE: Ohio
; COUNTRY: USA
; ZIP: 45215-6300
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,617
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/971,909
; FILING DATE: 18-DEC-1992
; APPLICATION NUMBER: US 07/557,288
; FILING DATE: 24-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US91/04658
; FILING DATE: 28-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Boudreaux, William R
; REGISTRATION NUMBER: 35,796
; REFERENCE/DOCKET NUMBER: M01556 US-A

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (513) 948-6566
TELEFAX: (513) 948-7961
TELEX: 214320
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-432-617-4

Query Match 62.1%; Score 18; DB 1; Length 6;
Best Local Similarity 75.0%; Pred. No. 3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGP 4
:|||
Db 3 TSGP 6

RESULT 15
US-09-274-642-19
; Sequence 19, Application US/09274642A
; Patent No. 6071729
; GENERAL INFORMATION:
; APPLICANT: Jeffries, Thomas W.
; APPLICANT: Shi, Nian-Qing
; TITLE OF INVENTION: Disruption of cytochrome c gene in xylose-fermenting yeast
; FILE REFERENCE: cytochrome c
; CURRENT APPLICATION NUMBER: US/09/274,642A
; CURRENT FILING DATE: 1999-03-23
; EARLIER APPLICATION NUMBER: 60/080,493
; EARLIER FILING DATE: 1998-04-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 19
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: conserved
; OTHER INFORMATION: sequence
US-09-274-642-19

Query Match 62.1%; Score 18; DB 3; Length 6;
Best Local Similarity 75.0%; Pred. No. 3e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPSL 6
|||:
Db 1 GPNL 4

Search completed: March 10, 2004, 15:16:40
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: March 10, 2004, 15:08:31 ; Search time 52 Seconds
(without alignments)
32.502 Million cell updates/sec

Title: US-09-848-834A-6

Perfect score: 29

Sequence: 1 SSGPSL 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 59163

Minimum DB seq length: 0
Maximum DB seq length: 6

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	6	5 AAU11417	AAU11417 Synthetic
2	21	72.4	4	4 AAB68641	AAB68641 Peptide 1
3	21	72.4	4	5 AAU11416	AAU11416 Synthetic
4	21	72.4	5	4 AAU00361	AAU00361 Binding m
5	21	72.4	5	5 ABG77493	ABG77493 Targettin
6	21	72.4	6	2 AAR56526	AAR56526 Conserved
7	21	72.4	6	2 AAR76622	AAR76622 Saccharom
8	21	72.4	6	2 AAW67574	AAW67574 Linker pe
9	21	72.4	6	2 ADA25166	ADA25166 Chimeric
10	21	72.4	6	7 ADC89655	ADC89655 Synthetic
11	19	65.5	5	2 AAW45758	AAW45758 Apoptotic
12	19	65.5	6	7 ADE55098	ADE55098 Corticotr
13	19	65.5	6	7 ADE51384	ADE51384 CRF2 non-
14	18	62.1	4	2 AAW03253	AAW03253 Peptidyl-
15	18	62.1	5	2 AAR67442	AAR67442 OKT3 huma
16	18	62.1	5	4 AAB81169	AAB81169 Additiona
17	18	62.1	5	4 AAB81170	AAB81170 Additiona
18	18	62.1	5	6 ADA000982	ADA000982 Mouse Rag
19	18	62.1	6	2 AAR67444	AAR67444 OKT3 huma
20	18	62.1	6	2 AAR67445	AAR67445 OKT3 huma
21	18	62.1	6	2 AAY23488	AAY23488 V beta 6
22	18	62.1	6	6 ABP74904	ABP74904 Proteome
23	18	62.1	6	6 ADE565100	ADE565100 Corticotr
24	18	62.1	6	7 ADE51386	ADE51386 CRF2 non-
25	17	58.6	4	2 AAR61120	AAR61120 Generic h

26	17	58.6	4	2 AAR49967	AAR49967 Human hip
27	17	58.6	4	4 AAB48424	AAB48424 Human ICA
28	17	58.6	4	6 ABJ36808	ABJ36808 G protein
29	17	58.6	5	2 AAR13676	AAR13676 Pentapept
30	17	58.6	5	2 AAR49969	AAR49969 Human hip
31	17	58.6	5	2 AAR49966	AAR49966 Human hip
32	17	58.6	5	2 AAR49968	AAR49968 Human hip
33	17	58.6	5	2 AAR49970	AAR49970 Human hip
34	17	58.6	5	2 AAR51540	AAR51540 Minotope
35	17	58.6	5	2 AAR55760	AAR55760 RNK Met-1
36	17	58.6	5	2 AAR55058	AAR55058 CD4 pepti
37	17	58.6	5	2 AAR87733	AAR87733 Spacer #2
38	17	58.6	5	2 AAR69908	AAR69908 Pentameri
39	17	58.6	5	2 AAR98654	AAR98654 Peptide 3
40	17	58.6	5	2 AAY16734	AAY16734 WO9914235
41	17	58.6	5	2 AAY22478	AAY22478 Human AUR
42	17	58.6	5	3 AAY83348	AAY83348 Peptide m
43	17	58.6	5	4 AAU00369	AAU00369 Binding m
44	17	58.6	5	5 ABC04532	ABJ04532 Molt-4 le
45	17	58.6	5	5 AAE13445	AAE13445 Synthetic

ALIGNMENTS

RESULT 1
AAU11417
ID AAU11417 standard; peptide; 6 AA.

XX AC AAU11417;

DT 12-MAR-2002 (first entry)

XX Synthetic spacer peptide #2.

XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer;
KW spacer peptide.

XX Synthetic.

XX PN WO200185763-A2.

XX PD 15-NOV-2001.

XX PF 04-MAY-2001; 2001WO-US014363.

XX PR 05-MAY-2000; 2000US-0202328P.

XX PA (APHT-) APHTON CORP.

XX PI Grimes S, Michaeli D, Stevens VC;

XX DR WPI; 2002-049440/06.

XX PT Novel synthetic immunogen for inducing immune response against

PT gonadotropin releasing hormone, comprises fusion peptide having

PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope

PT or its analog.

XX PS Claim 10; Page 6; 43pp; English.

XX The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotropin releasing hormone (GnRH) also known as
CC luteinising hormone releasing hormone (LHRH) comprising a fusion peptide
CC which comprises a promiscuous helper T-cell peptide epitope and
CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
CC useful inducing an immune response against GnRH in an animal subject, and
CC as such is useful as a contraceptive and in the treatment of diseases
CC such as cancer (of the breast, uterus and other gynaecological cancer),

CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC spacer peptide used in the immunogen of the invention

XX SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 5; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6
 |||||
 DB 1 SSGPSL 6

RESULT 2

AB68641
 ID AAB68641 standard; peptide; 4 AA.

XX AC AAB68641;

XX DT 30-APR-2001 (first entry)

XX DE Peptide linker.

XX KW Cytostatic; immune response; HER-2; human; epitope; cancer; breast;
 KW ovarian; lung; prostate; colon.

XX OS Unidentified.

PN W0200108636-A2.

XX XX

PD 08-FEB-2001.

XX PF 03-AUG-2000; 2000WO-US021222.

XX PR 03-AUG-1999; 99US-0146869P.

XX PA (OHIS) UNIV OHIO STATE.

XX PI Kaumaya PT, Stevens VC, Triozzi PL;

XX XX WPI; 2001-182849/18.

XX PT Compositions comprising polypeptides and polynucleotides for stimulating
 PT the immune system and for treating malignancies associated with
 PT overexpression of the HER-2 protein.

XX PS Claim 5; Page 38; 51pp; English.

XX CC The present invention relates to compositions for stimulating the immune
 CC system and for treating malignancies associated with overexpression of
 CC the HER-2 protein. The compositions comprise immunogenic groups of the
 CC HER-2 proteins. The present sequence is one such peptide used in the
 CC compositions of the present invention. The compositions can be used for
 CC treating cancer, e.g. breast, ovarian, lung, prostate and colon cancers

XX SQ Sequence 4 AA;

Query Match 72.4%; Score 21; DB 4; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPSL 6
 |||||
 DB 1 GPSL 4

RESULT 3

AAU11416
 ID AAU11416 standard; peptide; 4 AA.

XX XX

AC AAU11416;
 XX DT 12-MAR-2002 (first entry)
 XX DE Synthetic spacer peptide #1.

XX KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;
 KW luteinising hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer;
 KW spacer peptide.

XX OS Synthetic.

XX PN W0200185763-A2.

XX XX

PD 15-NOV-2001.

XX PF 04-MAY-2001; 2001WO-US014363.

XX PR 05-MAY-2000; 2000US-0202328P.

XX PA (APHT-) APHTON CORP.

XX XX Grimes S, Michaeli D, Stevens VC;

XX DR WPI; 2002-049440/06.

XX PT Novel synthetic immunogen for inducing immune response against
 PT gonadotropin releasing hormone, comprises fusion peptide having
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 PT or its analog.

XX PS Claim 10; Page 6; 43pp; English.

XX CC The invention relates to a synthetic immunogen for inducing specific
 CC antibodies against gonadotropin releasing hormone (GnRH) also known as
 CC luteinising hormone releasing hormone (LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC spacer peptide used in the immunogen of the invention

XX SQ Sequence 4 AA;

Query Match 72.4%; Score 21; DB 5; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPSL 6
 |||||
 DB 1 GPSL 4

RESULT 4

AAU00361
 ID AAU00361 standard; peptide; 5 AA.

XX AC AAU00361;

XX DT 04-JUL-2001 (first entry)

XX DE Binding motif #15 used in study of receptor binding motifs.

XX KW Common Beta chain; Beta-c; binding motif; cancer; therapeutic;
 KW acute myeloid leukaemia; AML; inflammatory disease; asthma;
 KW rheumatoid arthritis; cell proliferative disease.

```

XX OS Synthetic.
XX PN WO200119847-A1.
XX XX
XX PD 22-MAR-2001.
XX XX
XX PF 15-SEP-2000; 2000WO-AU001118.
XX XX
XX PR 15-SEP-1999; 99AU-00002875.
XX PR 12-JUL-2000; 2000AU-00008733.
XX XX
XX PA (MEDV-) MEDVET SCI PTY LTD.
XX PA (BERN/) BERNDT M C.
XX XX
XX PI Guthridge MA, Stomeki FC, Lopez AF;
XX DR WPI; 2001-244778/25.
XX XX
XX PT New binding motif of a receptor capable of binding to cytoplasmic
XX PT protein, for use as a tool for treating and preventing cell proliferative
XX PT diseases such as acute myeloid leukemia and cancer.
XX XX
XX PS Claim 8; Page 60; 101pp; English.
XX XX
XX CC The sequence represents the amino acid sequence of binding motif #15 used
XX CC in study of a binding motif (I) of a receptor capable of binding a
XX CC cytoplasmic protein. (I) comprises an amino acid sequence in which at
XX CC least one amino acid is serine/threonine. An antagonist to the
XX CC interaction of (I) to a cytoplasmic protein is useful as a cancer
XX CC therapeutic, especially for preventing or treating leukaemia such as
XX CC acute myeloid leukaemia (AML), and inflammatory diseases e.g., asthma and
XX CC rheumatoid arthritis, and for preventing functions related to cell
XX CC activation. The interaction of (I) and cytoplasmic protein is useful as a
XX CC tool for treating and preventing cell proliferative diseases such as AML
XX CC and cancer
XX XX
XX SQ Sequence 5 AA;

Query Match 72.4%; Score 21; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGP 4
Db |||||
2 SSGP 5

RESULT 5
ABG77493
XX ID ABG77493 standard; peptide; 5 AA.
XX AC ABG77493;
XX XX
XX DT 05-NOV-2002 (first entry)
XX DE
XX XX
XX KW Human; cytostatic; antiinflammatory; antidiabetic; cardiovascular;
XX KW immunomodulator; antibacterial; antiviral; gene therapy; cancer;
XX KW arthritis; diabetes; inflammatory disease; atherosclerosis;
XX KW autoimmune disease; bacterial infection; viral infection;
XX KW cardiovascular disease; degenerative disease.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO2000220723-A2.
XX XX
XX PD 14-MAR-2002.
XX XX
XX PF 07-SEP-2001; 2001WO-US028044.
XX XX
XX PR 08-SEP-2000; 2000US-0231266P.

17-JAN-2001; 2001US-00765101.
(TEXA ) UNIV TEXAS SYSTEM.
Arap W, Pasqualini R;
WPI; 2002-599247/64.
New targeting peptides identified by phage display, useful for treating a
disease state, e.g. cancer, diabetes, inflammatory disease,
PT atherosclerosis, autoimmune disease, bacterial or viral infection or
cardiovascular disease.
XX
XX Claim 16; Page 63; 269pp; English.
XX
XX The invention describes an isolated peptide of 100 amino acids or less in
size. The peptide is useful for treating a disease state, e.g. cancer,
arthritis, diabetes, inflammatory disease, atherosclerosis, autoimmune
disease, bacterial infection, viral infection, cardiovascular disease or
degenerative disease. This sequence represents a human targeting peptide
selective for human organs, tissues or cell types
XX
XX Sequence 5 AA;

Query Match 72.4%; Score 21; DB 5; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPSL 6
Db |||||
1 GPSL 4

RESULT 6
AAR56526
XX ID AAR56526 standard; protein; 6 AA.
XX AC AAR56526;
XX XX
XX DT 25-MAR-2003 (revised)
XX DT 17-MAR-1995 (first entry)
XX XX
XX DE Conserved region of yeast protein kinase (HRR25).
XX XX
XX KW Protein kinase; immunogen; antibody; protein-tyrosine kinase;
XX KW protein-serine/threonine kinase; recombination; repair; screening;
XX KW detection.
XX XX
XX OS Saccharomyces cerevisiae.
XX XX
XX PN WO9417189-A2.
XX XX
XX PD 04-AUG-1994.
XX XX
XX PF 21-JAN-1994; 94WO-US0000795.
XX XX
XX PR 21-JAN-1993; 93US-00008001.
XX XX
XX (SALK ) SALK INST BIOLOGICAL STUDIES.
XX PA
XX PI Hoekstra MF;
XX XX
XX DR WPI; 1994-264102/32.
XX XX
XX PT Wild-type and mutant protein kinase genes and encoded polypeptide(s) -
XX PT Useful in screening for compositions which may effect DNA double strand
XX PT break repair activity.
XX PS Disclosure; Page 95; 121pp; English.
XX XX
XX CC This sequence is a conserved region of the yeast protein kinase
XX CC designated HRR25. Nucleic acid probes may be synthesised based on this
XX CC sequence and then used to identify sequences encoding HRR25-like protein

```

CC kinases. Host cells stably transformed with the protein kinase encoding
 CC DNA may be used for the expression of the protein kinase such that the
 CC expressed protein is "displayed" on the host cell surface. The cells may
 CC then be used as immunogen for the production of antibodies. The host
 CC cells may also be used for the large scale production of the protein
 CC kinase. The expressed protein being either isolated from the cell surface
 CC or from the culture medium. Recombinant HRR25 like proteins display a
 CC number of properties which are unique among the eukaryotic protein
 CC kinases e.g. the HRR25 protein possesses both protein-tyrosine kinase and
 CC protein-serine/threonine kinase activities. Also, HRR25 operates to
 CC promote repair of DNA strand breaks at a specific nucleotide sequence and
 CC is the only protein kinase known to have such recombination/repair
 CC promoting activity. Recombinant HRR25-like proteins and host cells
 CC expressing them are useful in screening methods designed to examine the
 CC effects of various compositions on DNA break repair and protein kinase
 CC activities of the proteins. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 6 AA;

Query Match 72.4%; Score 21; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPSL 6
 DB 1 GPSL 4

RESULT 7
 AAR76622
 ID AAR76622 standard; peptide; 6 AA.

XX AC

XX AAR76622;

XX 25-JAN-1996 (first entry)

XX DE Saccharomyces cerevisiae protein-kinase HRR25 conserved motif.

XX Yeast; fungus; casein-kinase-I; HRR25; protein-tyrosine-kinase;
 KW protein-serine/threonine-kinase; enzyme; DNA repair; DNA recombination.
 XX OS Saccharomyces cerevisiae.

XX FH Key Location/Qualifiers
 FT Peptide 1..6

XX PN W09519993-A1.

XX PD 27-JUL-1995.

XX PF 23-JAN-1995; 95WO-US000955.

XX PR 21-JAN-1994; 94US-00185359.

XX PA (SALK) SALK INST BIOLOGICAL STUDIES.

XX PI Hoekstra MF;

XX DR WPI; 1995-269419/35.

XX New monoclonal antibodies against human casein Kinase class I enzymes -
 FT useful for purificn. and determ. of these enzymes and to modulate their
 PT receptor -ligand binding, also new hybridomas.

XX PS Disclosure; Page 98; 125pp; English.

XX CC This sequence representing a conserved region of AA residues in HRR25
 CC which can be used to identify HRR25-like genes from other spp. This
 CC conserved motif can be used as a probe for identification and isolation
 CC of HRR25-like genes (homologs)

XX SQ Sequence 6 AA;

Query Match 72.4%; Score 21; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPSL 6
 DB 1 GPSL 4

RESULT 8
 AAW67574
 ID AAW67574 standard; peptide; 6 AA.

XX AC

XX AAW67574;

XX DT 02-MAR-1999 (first entry)

XX DE Linker peptide for fimbria/T-cell epitope chimeric peptide.

XX KW Chimeric; non-typable Haemophilus influenzae; fimbria; T-cell epitope;
 KW immunogenic composition; immune response.

XX OS Synthetic.

XX PN US5843464-A.

XX PD 01-DEC-1998.

XX PF 02-JUN-1995; 95US-00460502.

XX PR 02-JUN-1995; 95US-00460502.

XX PA (OHIS) UNIV OHIO STATE.

XX PI Kaumaya FTP, Bakaletz IO;

XX WPI; 1999-044514/04.

XX Synthetic chimeric fimbria peptide - useful for vaccination against non-
 FT typable Haemophilus influenzae.

XX PS Disclosure; Col 4; 16pp; English.

XX The invention relates to the manufacture of a synthetic chimeric peptide
 CC comprising a non-typable Haemophilus influenzae fimbria peptide fused via
 CC a linker peptide to a T-cell epitope peptide. The chimeric peptide is
 CC used in immunogenic compositions which induce an immune response against
 CC non-typable Haemophilus influenzae. This sequence represents an example
 CC of the linker sequence used to generate the chimeric peptide

XX SQ Sequence 6 AA;

Query Match 72.4%; Score 21; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPSL 6
 DB 1 GPSL 4

RESULT 9
 ADA25166
 ID ADA25166 standard; peptide; 6 AA.

XX AC

XX ADA25166;

XX DT 20-NOV-2003 (first entry)

XX DE Chimeric fimbria peptide, peptide linker.

XX KW fimbria; non-typable Haemophilus influenzae; NTHi infection;
 XX otitis media.

XX OS Synthetic.

XX .PN US6436405-B1.

XX 20-AUG-2002.

XX 04-SEP-1998; 98US-00148711.

XX 02-JUN-1995; 95US-00460502.

XX (CHIS) UNIV OHIO STATE.

XX Bakaletz LO, Kaumaya PTP;

XX WPI; 2003-615247/58.

XX Synthetic chimeric fimbrin peptide, useful for treating Haemophilus influenzae infections.

XX Disclosure; Col 4; 16pp; English.

XX The invention relates to a synthetic chimeric fimbrin peptide. The peptide is useful for treating a non-typable Haemophilus influenzae (NTHi) infection and otitis media. The synthetic peptides do not require tedious purification techniques. The present sequence represents the amino acid sequence of a chimeric fimbrin peptide, peptide linker.

XX Sequence 6 AA;

Query Match 72.4%; Score 21; DB 6; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 GPSL 6
|||

Db 1 GPSL 4

RESULT 10

ID ADC89655 standard; peptide; 6 AA.

AC ADC89655;

DT 01-JAN-2004 (first entry)

DE Synthetic linker peptide.

XX Fimbrin; T cell epitope; vaccine; otitis media; auditory;

XX antiinflammatory; linker.

XX Synthetic.

XX US2003113344-A1.

XX 19-JUN-2003.

XX 19-AUG-2002; 2002US-00223711.

XX 04-SEP-1998; 98US-00148711.

XX (BAKA/) BAKALETZ L O.

XX (KAUM/) KAUMAYA P T P.

XX Bakaletz LO, Kaumaya PTP;

XX WPI; 2003-810881/76.

XX Novel synthetic chimeric fimbrin peptide LB1 or LB2 comprising a first peptide unit, T cell epitope as second peptide unit and third linker peptide unit, useful for preventing or reducing severity of otitis media.

PS Disclosure; SEQ ID NO 4; 15pp; English.

XX The invention relates to a synthetic chimaeric fimbrin peptide LB1 or LB2 comprises a first peptide unit derived from H. influenzae fimbrin , a second peptide unit containing a T cell epitope and a third linker peptide which connects the first peptide to the second. The chimaeric peptide is useful for inducing an immune response in animals against non-typable Haemophilus influenzae (NTHi) and for preventing or reducing adherence of NTHi to host cells thereby preventing or reducing the severity of otitis media. The present sequence is a synthetic linker peptide for use in the chimaeric peptides of the invention.

XX Sequence 6 AA;

Query Match 72.4%; Score 21; DB 7; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 GPSL 6
|||

Db 1 GPSL 4

RESULT 11

AAW45758

ID AAW45758 standard; peptide; 5 AA.

XX AAW45758;

XX 19-JUN-1998 (first entry)

XX Apoptotic protease activating factor-2 tryptic peptide #7.

XX Cytochrome c; tryptic peptide; Apaf-2; HeLa cell S-100; apoptosis;

XX apoptotic protease activating factor-2; cytosol; antitumour.

XX Homo sapiens.

XX WO9802579-A1.

XX 22-JAN-1998.

XX 11-JUL-1997; 97WO-US012090.

XX 12-JUL-1996; 96US-0021268P.

XX (UYEM-) UNIV EMORY.

XX Wang X, Liu X;

XX WPI; 1998-110623/10.

XX Cell-free assay to identify compositions that regulate apoptosis - from their effect on e.g. cytosolic cytochrome c levels in cellular extracts, particularly to identify antitumour agents.

XX Example 8; Page 22; 42pp; English.

XX This sequence represents an apoptotic protease activating factor-2 (Apaf-2) tryptic peptide which corresponds to residues 28-32 of human cytochrome c. The invention relates to an in vitro model for studying the regulation of apoptosis. Compositions are assayed in vitro for regulation of apoptosis by (a) preparing a 10000 g supernatant extract (A) from cells, not committed to apoptosis, from a multicellular eukaryote; (b) adding test composition to (A); (c) preparing control assays comprising (A) and a known inhibitor or known inducer of apoptosis, or a compound known to have no effect on apoptosis; (d) assessing either (i) activation of apoptosis from an increase in cytosolic cytochrome c; CPP32 proase activity or ability to fragment genomic DNA, relative to system without added test compounds or (ii) inhibition of apoptosis by a reduction in these quantities in a mixture containing a known inducer, relative to control without test compound. The test is performed on mammalian cells, specifically HeLa cells. Cytochrome c is determined by immunoassay and

CC CPP32 by adding radiolabelled poly(adenosine diphosphate-ribose)
 CC polymerase (PAP) or labelled sterol regulatory binding protein (SRBP).
 CC and then detecting fragments of these compounds by sodium dodecylsulphate
 CC -polyacrylamide gel electrophoresis (SDS-PAGE). DNA fragmentation is
 CC determined by adding intact mammalian cell nuclei, incubating, extracting
 CC genomic DNA and analysing its size distribution. The extract contains
 CC enough dATP and/or corresponding diphosphate (dADP) to allow induction of
 CC an apoptotic response. dATP and soluble cytochrome c are required for
 CC initiation of apoptosis in the cell-free system. The method is especially
 CC used to identify chemotherapeutic agents that can induce apoptosis in
 CC tumour cells even when these express the Bcl-2 protein associated with
 CC resistance to chemotherapy. Compounds identified this way may also (a)
 CC improve the effect of apoptosis-inducing chemotherapeutic agents or (b)
 CC antagonise deoxyadenosine triphosphate (dATP) in the cytosol of patients
 CC with adenosine deaminase deficiency (severe combined immune deficiency),
 CC as a potential treatment for this condition. Unlike known methods, this
 CC process does not use cells in which the apoptotic pathway has already
 CC been induced, i.e. it is not subject to interference from other apoptosis
 CC -inducing factors or other conditions
 XX
 SQ Sequence 5 AA;

Query Match 65.5%; Score 19; DB 2; Length 5;
 Best Local Similarity 60.0%; Pred. No. 1.4e+06;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 SGPSL 6
 :||:
 Db 1 TGNL 5

RESULT 12

ADE65098
 ID ADE65098 standard; peptide; 6 AA.

AC ADE65098;

DT 29-JAN-2004 (first entry)

DE Corticotropin-releasing factor-2 polypeptide, SEQ ID NO 401.

XX corticotropin-releasing factor-2; CRF2; myopathic; osteopathic;
 KW hypotensive; cardiant; vasotropic; antimigraine; cerebroprotective;
 KW neurotropic; neuroprotective; anorectic; antidiabetic; analgesic;
 KW antiallergic; tranquilizer; anxiolytic; antidepressant; antiarthritic;
 KW gene therapy.

XX Unidentified.

OS WO2003062277-A1.

PN 31-JUL-2003.

PD 16-JAN-2003; 2003WO-US001454.

PF 16-JAN-2002; 2002US-0349117P.

PR 29-APR-2002; 2002US-0376337P.

PR 14-JUN-2002; 2002US-038895P.

PR 19-SEP-2002; 2002US-0411988P.

XX (PROC) PROCTER & GAMBLE CO.

XX Isfort RJ, Mazur WA;

XX WPI; 2003-787975/74.

XX New non-native peptide derived from corticotropin-releasing factor-2,

XX useful for treatment and prevention of e.g. muscular atrophy, also

XX related nucleic acid and antibodies.

XX Example 2; SEQ ID NO 401; 304pp; English.

XX The invention relates to a novel non-native peptide derived from

CC corticotropin-releasing factor-2 (CRF2). The CRF2 peptides have the
 CC following activities: myopathic, osteopathic, hypotensive, cardiant,
 CC vasotropic, antimigraine, cerebroprotective, neurotropic, neuroprotective,
 CC anorectic, antidiabetic, analgesic, antiallergic, tranquilizer,
 CC anxiolytic, antidepressant and antiarthritic. The CRF2 peptides, and
 CC related compounds derived from other proteins, are used to prevent or
 CC treat disorders modulated by the CRF2 receptor, e.g. skeletal muscle
 CC atrophy or wasting, and bone disorders, however caused; heart/circulatory
 CC diseases (e.g. hypertension, congestive heart failure, heart attack,
 CC reperfusion injury, migraine, stroke, memory loss, Alzheimer's disease,
 CC dementia); joint disorders (osteoarthritis or rheumatoid arthritis);
 CC metabolic disease (obesity or diabetes); pain; allergy; stress; anxiety;
 CC low levels of adrenocorticotrophic hormone; anorexia nervosa; depression;
 CC also to reduce body temperature and to control appetite or cognitive
 CC function. Nucleic acids, optionally labelled, that encode the CRF2
 CC peptides are used as primers and probes for amplification, also for gene
 CC synthesis and for recombinant production of CRF2 peptides, including use
 CC in gene therapy. Antibodies specific for the CRF2 peptides are used to
 CC evaluate expression of the CRF2 peptides after gene therapy. This
 CC sequence represents a novel native CRF polypeptide of the invention.

XX SQ Sequence 6 AA;

Query Match 65.5%; Score 19; DB 7; Length 6;
 Best Local Similarity 66.7%; Pred. No. 1.4e+06;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSGPSL 6

Db 1 SDNPFL 6

RESULT 13

ADE51384
 ID ADE51384 standard; peptide; 6 AA.

AC ADE51384;

DT 29-JAN-2004 (first entry)

DE CRF2 non-native polypeptide, SEQ ID NO 401.

XX non-native; corticotropin-releasing factor-2; CRF2; myopathic;
 KW osteopathic; hypotensive; cardiant; vasotropic; antimigraine;
 KW cerebroprotective; neurotropic; neuroprotective; anorectic; antidiabetic;
 KW analgesic; antiallergic; tranquilizer; anxiolytic; antidepressant;
 KW antiarthritic.

XX Unidentified.

OS WO2003062268-A2.

PN 31-JUL-2003.

PD 16-JAN-2003; 2003WO-US001451.

PF 16-JAN-2002; 2002US-0349117P.

PR 29-APR-2002; 2002US-0376337P.

PR 14-JUN-2002; 2002US-038895P.

PR 19-SEP-2002; 2002US-0411988P.

XX (PROC) PROCTER & GAMBLE CO.

XX Isfort RJ, Mazur WA;

XX WPI; 2003-787974/74.

XX New non-native peptide derived from corticotropin-releasing factor-2,

XX useful for treatment and prevention of e.g. muscular atrophy, also

XX related nucleic acid and antibodies.

XX Example 2; SEQ ID NO 401; 300pp; English.

Best Local Similarity 75.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SGPS 5
: |||
Db 2 AGPS 5

Search completed: March 10, 2004, 15:14:17
Job time : 53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 15:15:37 ; Search time 33 Seconds
(without alignments)
38.392 Million cell updates/sec

Title: US-09-848-834A-6

Perfect score: 29

Sequence: 1 SSGPSL 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 21259

Minimum DB seq length: 0

Maximum DB seq length: 6

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	6	9	US-09-848-834A-6
2	21	72.4	4	9	US-09-848-834A-5
3	21	72.4	4	14	US-10-285-976-224
4	21	72.4	5	13	US-10-099-895-18
5	21	72.4	6	14	US-10-223-711-4
6	19	65.5	5	8	US-08-831-525-7
7	19	65.5	5	14	US-10-198-590-7
8	19	65.5	6	14	US-10-315-964A-401
9	19	65.5	6	14	US-10-317-251A-401
10	19	65.5	6	14	US-10-317-252A-401
11	18	62.1	5	14	US-10-112-488-69
12	18	62.1	5	14	US-10-112-488-70
13	18	62.1	5	14	US-10-267-286A-20
14	18	62.1	5	14	US-10-197-927-16
15	18	62.1	6	9	US-09-741-148A-15

16	18	62.1	6	14	US-10-354-577-15	Sequence 15, Appl
17	18	62.1	6	14	US-10-267-286A-22	Sequence 22, Appl
18	18	62.1	6	14	US-10-267-286A-23	Sequence 23, Appl
19	18	62.1	6	14	US-10-315-964A-403	Sequence 403, App
20	18	62.1	6	14	US-10-317-251A-403	Sequence 403, App
21	18	62.1	6	14	US-10-317-252A-403	Sequence 403, App
22	18	62.1	6	15	US-10-394-980-187	Sequence 187, App
23	17	59.6	4	10	US-09-852-910-158	Sequence 158, App
24	17	59.6	4	15	US-10-411-336A-158	Sequence 158, App
25	17	59.6	5	9	US-09-012-135A-6	Sequence 6, Appli
26	17	59.6	5	13	US-10-099-895-26	Sequence 26, Appl
27	17	59.6	5	14	US-10-153-334-44	Sequence 44, Appl
28	17	59.6	5	14	US-10-029-065-22	Sequence 22, Appl
29	17	59.6	5	15	US-10-137-867-194	Sequence 194, App
30	17	59.6	6	8	US-08-860-844-102	Sequence 102, App
31	17	59.6	6	14	US-10-043-487-450	Sequence 450, App
32	17	59.6	6	14	US-10-267-286A-18	Sequence 18, Appl
33	17	59.6	6	14	US-10-267-286A-19	Sequence 19, Appl
34	17	59.6	6	14	US-10-267-286A-21	Sequence 21, Appl
35	17	59.6	6	14	US-10-407-543-102	Sequence 102, App
36	17	59.6	6	15	US-10-297-337-1	Sequence 1, Appli
37	16	55.2	5	14	US-10-206-699-61	Sequence 61, Appl
38	16	55.2	5	14	US-10-206-699-198	Sequence 198, App
39	16	55.2	6	10	US-09-518-737-6	Sequence 6, Appli
40	16	55.2	6	14	US-10-262-435-11	Sequence 11, Appl
41	16	55.2	6	14	US-10-262-435-97	Sequence 97, Appl
42	16	55.2	6	14	US-10-348-232-242	Sequence 242, App
43	16	55.2	6	14	US-10-086-208-11	Sequence 11, Appl
44	16	55.2	6	14	US-10-086-208-97	Sequence 97, Appl
45	16	55.2	6	15	US-10-156-214A-281	Sequence 281, App

ALIGNMENTS

RESULT 1
US-09-848-834A-6
; Sequence 6, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-09-848-834A-6

Query Match 100.0%; Score 29; DB 9; Length 6;
Best Local Similarity 100.0%; Pred.No. 7.le+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSGPSL 6

Db 1 SSGPSL 6

RESULT 2

US-09-848-834A-5
; Sequence 5, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens

```

RESULT 4
US-10-099-895-18
; Sequence 18, Application US/10099895
; Publication No. US20020177166A1
; GENERAL INFORMATION:
; APPLICANT: BERNDT, Michael C
; APPLICANT: STOMSKI, Frank C
; APPLICANT: LOPEZ, Angel F

```

NAME: Yoo-Warren, Heeja
REGISTRATION NUMBER: 45,495
REFERENCE/DOCKET NUMBER: 45-96A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-198-590-7

Query Match 65.5%; Score 19; DB 14; Length 5;
Best Local Similarity 60.0%; Pred. No. 7.1e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SGPSL 6
Db 1 TGPNL 5

RESULT 8
US-10-315-964A-401
; Sequence 401, Application US/10315964A
; Publication No. US20030148956A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Isfort, Robert J
; APPLICANT: Mazur, Wieslaw A
; TITLE OF INVENTION: Corticotropin Releasing Factor 2 Receptor Agonists
; FILE REFERENCE: 8847M3
; CURRENT APPLICATION NUMBER: US/10/315,964A
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/349,117
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: US 60/376,337
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 60/388,895
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/411,988
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 530
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 401
; LENGTH: 6
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Chemically synthesized artificial peptide
US-10-315-964A-401

Query Match 65.5%; Score 19; DB 14; Length 6;
Best Local Similarity 66.7%; Pred. No. 7.1e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SSGPSL 6
Db 1 SDNPSL 6

RESULT 9
US-10-317-251A-401
; Sequence 401, Application US/10317251A
; Publication No. US20030148957A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Isfort, Robert J

COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/891,525
FILING DATE: 11-JUL-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/021,268
FILING DATE: 12-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 45-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
US-08-891-525-7

Query Match 65.5%; Score 19; DB 8; Length 5;
Best Local Similarity 60.0%; Pred. No. 7.1e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SGPSL 6
Db 1 TGPNL 5

RESULT 7
US-10-198-590-7
; Sequence 7, Application US/10198590
; Publication No. US20030032045A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; Liu, Xuesong
; TITLE OF INVENTION: Regulation of Apoptosis and In Vitro
; Model for Studies Thereof
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/198,590
FILING DATE: 18-JUL-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/891,525
FILING DATE: 11-JUL-1997
APPLICATION NUMBER: US 60/021,268
FILING DATE: 12-JUL-1996
ATTORNEY/AGENT INFORMATION:

```
; APPLICANT: Mazur, Wieslaw A
; TITLE OF INVENTION: Corticotropin Releasing Factor 2 Receptor Agonists
; FILE REFERENCE: 8847M2
; CURRENT APPLICATION NUMBER: US/10/317,251A
; CURRENT FILING DATE: 2002-12-11
; PRIOR APPLICATION NUMBER: US 60/349,117
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: US 60/376,337
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 60/388,895
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/411,988
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 530
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 401
; LENGTH: 6
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Chemically synthesized artificial peptide
US-10-317-251A-401

Query Match          65.5%; Score 19; DB 14; Length 6;
Best Local Similarity 86.7%; Pred. No. 7.1e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 SSGPSL 6
Db      1 SDNPSL 6

RESULT 10
US-10-317-252A-401
; Sequence 401, Application US/10317252A
; Publication No. US20030148958A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Isfort, Robert J
; APPLICANT: Mazur, Wieslaw A
; TITLE OF INVENTION: Corticotropin Releasing Factor 2 Receptor Agonists
; FILE REFERENCE: 8847M
; CURRENT APPLICATION NUMBER: US/10/317,252A
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/349,117
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: US 60/376,337
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 60/388,895
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/411,988
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 530
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 401
; LENGTH: 6
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Chemically synthesized artificial peptide
US-10-317-252A-401

Query Match          65.5%; Score 19; DB 14; Length 6;
Best Local Similarity 86.7%; Pred. No. 7.1e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 SSGPSL 6
Db      1 SDNPSL 6

RESULT 11
US-10-112-488-69
```

```
; Sequence 69, Application US/10112488
; Publication No. US20030082746A1
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, Yoshihimi
; APPLICANT: DATE, Masayo
; APPLICANT: UMEZAWA, Yukiko
; APPLICANT: YOKOYAMA, Keiichi
; APPLICANT: MATSUI, Hiroshi
; TITLE OF INVENTION: PROCESS FOR PRODUCING TRANSGLUTAMINASE
; FILE REFERENCE: 219286USOCONT
; CURRENT APPLICATION NUMBER: US/10/112,488
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: PCT/JP00/06780
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: JP2000-280098
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: JP11-280098
; PRIOR FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 69
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-112-488-69

Query Match          62.1%; Score 18; DB 14; Length 5;
Best Local Similarity 75.0%; Pred. No. 7.1e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SSGP 4
Db      1 SAGP 4

RESULT 12
US-10-112-488-70
; Sequence 70, Application US/10112488
; Publication No. US20030082746A1
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, Yoshihimi
; APPLICANT: DATE, Masayo
; APPLICANT: UMEZAWA, Yukiko
; APPLICANT: YOKOYAMA, Keiichi
; APPLICANT: MATSUI, Hiroshi
; TITLE OF INVENTION: PROCESS FOR PRODUCING TRANSGLUTAMINASE
; FILE REFERENCE: 219286USOCONT
; CURRENT APPLICATION NUMBER: US/10/112,488
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: PCT/JP00/06780
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: JP2000-280098
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: JP11-280098
; PRIOR FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 70
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-112-488-70

Query Match          62.1%; Score 18; DB 14; Length 5;
Best Local Similarity 75.0%; Pred. No. 7.1e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SSGP 4
Db      1 SAGP 4
```

Db 1 SACP 4

RESULT 13

US-10-267-286A-20
; Sequence 20, Application US/10267286A
; Publication No. US20030108548A1
; GENERAL INFORMATION:
; APPLICANT: BLUESTONE, JEFFREY A.
; APPLICANT: ZIVIN, ROBERT A.
; APPLICANT: JOLLIFFE, LINDA K.
; TITLE OF INVENTION: METHODS AND MATERIALS FOR MODULATION OF THE
; TITLE OF INVENTION: IMMUNOSUPPRESSIVE ACTIVITY AND TOXICITY OF MONOCLONAL
; TITLE OF INVENTION: ANTIBODIES
; FILE REFERENCE: TOLT-004USC1
; CURRENT APPLICATION NUMBER: US/10/267,286A
; PRIOR FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: 08/557,050
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: PCT/US94/06198
; PRIOR FILING DATE: 1994-06-01
; PRIOR APPLICATION NUMBER: 08/070,116
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 20
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-267-286A-20

Query Match

Best Local Similarity 62.1%; Score 18; DB 14; Length 5;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SGPS 5

Db 2 AGPS 5

RESULT 14

US-10-197-927-16
; Sequence 16, Application US/10197927
; Publication No. US20030166138A1
; GENERAL INFORMATION:
; APPLICANT: Kinsella, Todd
; APPLICANT: Ohashi, Cara
; APPLICANT: Anderson, Dave
; TITLE OF INVENTION: Cyclic Peptides and Analogs Useful to Treat Allergies
; FILE REFERENCE: RIGL-002/01US
; CURRENT APPLICATION NUMBER: US/10/197,927
; CURRENT FILING DATE: 2003-01-23
; PRIOR APPLICATION NUMBER: 60/358,827
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 16
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically synthesized cyclic peptide
US-10-197-927-16

Query Match

Best Local Similarity 62.1%; Score 18; DB 14; Length 5;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGP 4

|:|

Db 1 STGP 4

RESULT 15

US-09-741-148A-15
; Sequence 15, Application US/09741148A
; Patent No. US20020076750A1
; GENERAL INFORMATION:
; APPLICANT: Chunhua YAN et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO00566
; CURRENT APPLICATION NUMBER: US/09/741,148A
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: 60/206,982
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-741-148A-15

Query Match 62.1%; Score 18; DB 9; Length 6;
Best Local Similarity 75.0%; Pred. No. 7.1e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SGPS 5

Db 2 TGPS 5

Search completed: March 10, 2004, 15:20:50
Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 15:12:07 ; Search time 39 Seconds
(without alignments)
48.541 Million cell updates/sec

Title: US-09-848-834A-6

Perfect score: 29

Sequence: 1 SSGPSL 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 11

Minimum DB seq length: 0

Maximum DB seq length: 6

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	8	27.6	6	5 P83569	sepia offic
2	8	27.6	6	10 P82341	spinacia ol
3	7	24.1	5	13 P83308	gallus gall
4	6	20.7	2	5 P83570	sepia offic
5	4	13.8	4	5 P83568	sepia offic
6	4	13.8	4	11 Q08433	rattus sp.
7	4	13.8	6	10 P82181	spinacia ol
8	4	13.8	6	10 P82182	spinacia ol
9	2	6.9	5	2 P83073	bacillus ce
10	2	6.9	5	10 Q99007	hordeum vul
11	2	6.9	6	2 P83533	lactobacill

ALIGNMENTS

RESULT 1

P83569 PRELIMINARY; PRT; 6 AA.

AC P83569; (TREMELrel. 24, Created)

DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)

DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)

DE Sperm attracting peptide SepsAP.

OS Sepia officinalis (Common cuttlefish).

OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;

OC Decapodiformes; Sepioidea; Sepiidae; Sepia.

OX NCBI_TaxID=6610;

RN [1]

RP SEQUENCE, FUNCTION, DEVELOPMENTAL STAGE, MASS SPECTROMETRY, AND

RP AMIDATION.

RC TISSUE=egg;

RX PubMed=12207899;

RA Zatylny C., Marvin L., Gagnon J., Henry J.;

RT "Fertilization in Sepia officinalis: the first mollusk sperm-

RT attracting peptide.";

RL Biochem. Biophys. Res. Commun. 296:1186-1193(2002).

CC -!- FUNCTION: ATTRACTS SPERM INCREASING THE CHANCES OF GAMETE

CC -!- COLLISION.

CC -!- DEVELOPMENTAL STAGE: FIRST APPEARS IN THE OVARIAN FOLLICLES FURING

CC VITELLOGENESIS. ACCUMULATES IN THE OOCYTES BEFORE BEING SECRETED

CC DURING FERTILIZATION. EXPRESSION CONTINUES IN THE EMBEDDED OOCYTE.

CC ACCUMULATES IN THE EGG CAPSULE AFTER FERTILIZATION.

CC -!- MASS SPECTROMETRY: MW=596.6; METHOD=MALDI.

KW Amidation.

FT MOD_RES 6 6 AMIDATION.

SQ SEQUENCE 6 AA; 597 MW; 72C8676AA0470000 CRC64;

Query Match 27.6%; Score 8; DB 5; Length 6;

Best Local Similarity 33.3%; Pred. No. 1e+06;

Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PSL 6

Db 4 PGV 6

RESULT 2

P82341 PRELIMINARY; PRT; 6 AA.

ID P82341 (TREMELrel. 15, Created)

AC P82341; (TREMELrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMELrel. 25, Last annotation update)

DE Chloroplast 30S ribosomal protein S19 beta (Fragment).

OS Spinacia oleracea (Spinach).

OG Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Caryophyllales; Amaranthaceae; Spinacia.

OX NCBI_TaxID=3562;

RN [1]

RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.

RC STRAIN=cv. ALVARO; TISSUE=leaf;

RX MEDLINE=20435797; PubMed=10874039;

RA Yamaguchi K., von Knoblauch K., Subramanian A.R.;

RT "The plastid ribosomal proteins. Identification of all the proteins in

RT the small subunit of an organelle ribosome (chloroplast).";

RL J. Biol. Chem. 37:28455-28465(2000).

CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.

CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.

CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.

CC -!- MASS SPECTROMETRY: MW=10477.0; METHOD=ELECTROSPRAY.

CC -!- MASS SPECTROMETRY: MW=10495; METHOD=MALDI.

CC -!- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PI. S19 BETA

CC FORM IS THE MINOR BASIC FORM.

CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 12 KDA.

CC -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.

DR GO; GO:0009507; C:chloroplast; IEA.

DR GO: 0019843; F:RNA binding; IEA.
 DR GO: 0003735; F:Structural constituent of ribosome; IEA.
 DR InterPro: IPR002222; Ribosomal_S19.
 DR PROSITE: PS00323; Ribosomal_S19; PARTIAL.
 KW Ribosomal protein; Chloroplast; rRNA-binding.
 FT NON_TER 6
 SQ SEQUENCE 6 AA; 732 MW; 63333735AA11C000 CRC64;
 Query Match 27.6%; Score 8; DB 10; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1e+06; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 SL 6
 Db 3 SL 4
 RESULT 3
 P83308 PRELIMINARY; PRT; 5 AA.
 AC P83308;
 DT 01-JUN-2002 (TREMELrel. 21, Created)
 DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE FMRamide-like neuropeptide (LPLRP-amide).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RC TISSUE=Brain;
 RX PubMed=6137771;
 RA Dockray G.J.; Reeve J.R. Jr., Shively J., Gayton R.J., Barnard C.S.;
 RT "A novel active pentapeptide from chicken brain identified by
 RT antibodies to FMRamide";
 RL Nature 305:328-330(1983).
 CC -1- FUNCTION: MAY FUNCTION AS A NEUROTRANSMITTER OR MODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRAMIDE RELATED PEPTIDE)
 CC FAMILY.
 DR GO: 0007218; P:neuropeptide signaling pathway; TAS.
 DR Neuropeptide; Amidation.
 FT MOD_RES 5
 SQ SEQUENCE 5 AA; 645 MW; 69D407376400000 CRC64;
 Query Match 24.1%; Score 7; DB 13; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1e+06; Indels 0; Gaps 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 P 4
 Db 2 P 2
 RESULT 4
 P83570 PRELIMINARY; PRT; 2 AA.
 AC P83570;
 DT 01-JUN-2003 (TREMELrel. 24, Created)
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Neuropeptide GWA.
 OS Sepia officinalis (Common cuttlefish).
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
 OC Decapodiformes; Sepioidea; Sepiidae; Sepia.
 OX NCBI_TaxID=6610;
 RN [1]
 RP SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND AMIDATION.
 RC TISSUE=Optic lobe;
 RX PubMed=9437704;
 RA Henry J., Favrel P., Boucaud-Camou E.;
 RT "Isolation and identification of a novel Ala-Pro-Gly-Trp-amide-related

RT peptide inhibiting the motility of the mature oviduct in the
 RT cuttlefish, Sepia officinalis.";
 RL Peptides 18:1469-1474(1997).
 CC -1- FUNCTION: REGULATORY NEUROPEPTIDE WITH MYOTROPIC ACTIVITY
 CC TARGETING THE DISTAL OVIDUCT. INHIBITS THE MOTILITY OF THE OVIDUCT
 CC BY DECREASING TONUS, FREQUENCY AND AMPLITUDE OF CONTRACTIONS.
 CC -1- MASS SPECTROMETRY: MW=259.9; METHOD=MALDI.
 DR GO: 0007218; P:neuropeptide signaling pathway; IEA.
 DR Neuropeptide; Amidation.
 FT MOD_RES 2
 SQ SEQUENCE 2 AA; 261 MW; 7378100000000000 CRC64;
 Query Match 20.7%; Score 6; DB 5; Length 2;
 Best Local Similarity 100.0%; Pred. No. 1e+06; Indels 0; Gaps 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 G 3
 Db 1 G 1

RESULT 5
 P83568 PRELIMINARY; PRT; 4 AA.
 AC P83568;
 DT 01-JUN-2003 (TREMELrel. 24, Created)
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Pheromone peptide ILME.
 OS Sepia officinalis (Common cuttlefish).
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
 OC Decapodiformes; Sepioidea; Sepiidae; Sepia.
 OX NCBI_TaxID=6610;
 RN [1]
 RP SEQUENCE, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND MASS
 RP SPECTROMETRY.
 RC TISSUE=Egg;
 RX PubMed=10944467;
 RA Zatylny C., Gagnon J., Boucaud-Camou E., Henry J.;
 RT "ILME: a waterborne pheromonal peptide released by the eggs of Sepia
 RT officinalis";
 RL Biochem. Biophys. Res. Commun. 275:217-222(2000).
 RN [2]
 RP SEQUENCE.
 RC TISSUE=Egg;
 RX PubMed=12207899;
 RA Zatylny C., Marvin L., Gagnon J., Henry J.;
 RT "Fertilization in Sepia officinalis: the first mollusk sperm-
 RT attracting peptide";
 RL Biochem. Biophys. Res. Commun. 296:1186-1193(2002).
 CC -1- FUNCTION: HAS MYOTROPIC ACTIVITY TARGETING THE GENITAL TRACT.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: FOLLICLE, FULLY GROWN OOCYTE AND EGG(EC2).
 CC -1- MASS SPECTROMETRY: MW=505.4; METHOD=MALDI.
 DR GO: 0005186; F:pheromone activity; IEA.
 KW Pheromone.
 SQ SEQUENCE 4 AA; 505 MW; 6B16972030000000 CRC64;

Query Match 13.8%; Score 4; DB 5; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1e+06;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 L 6
 Db 2 L 2

RESULT 6
 Q08433 PRELIMINARY; PRT; 4 AA.
 AC Q08433;
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)

01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DT Bilirubin UDP-glucuronosyltransferase (Fragment).
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10118;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Gunn;
 RX MEDLINE=91282758; PubMed=1840486;
 RA Sato H., Aono S., Kashiwazawa S., Koiwai O.;
 RT "Genetic defect of Bilirubin UDP-glucuronosyltransferase in the
 RL hyperbilirubinemic Gunn rat.";
 RL Biochem. Biophys. Res. Commun. 177:1161-1164 (1991).
 DR EMBL; S38636; AAB19259.1; -
 DR GO; GO:0016740; P:transferase activity; IEA.
 KW Transferase.
 FT NON TER
 SQ SEQUENCE 1 1
 SEQUENCE 4 AA; 473 MW; 633732C420000000 CRC64;
 Query Match 13.8%; Score 4; DB 11; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1e+06;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 L 6
 DB 3 L 3

RESULT 7
 ID P82181 PRELIMINARY; PRT; 6 AA.
 AC P82181
 DT 01-JUN-2000 (TrEMBLrel. 14, Created)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Chloroplast 50S ribosomal protein L10 beta (Fragment).
 OS Spinacia oleracea (Spinach).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllales; Amaranthaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=cv. ALVARO; TISSUE=Leaf;
 RX MEDLINE=20435798; PubMed=10874046;
 RA Yamaguchi K., Subramanian A.R.;
 RT "The plastid ribosomal proteins. Identification of all the proteins in
 the 50 S subunit of an organelle ribosome (chloroplast).";
 RL J. Biol. Chem. 275:28466-28482 (2000).
 CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
 CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
 CC -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
 DR GO; GO:0009507; C:chloroplast; IEA.
 DR GO; GO:0019843; F:rRNA binding; IEA.
 DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
 DR InterPro; IPR002363; Ribosomal L10eub.
 DR PROSITE; PS01109; RIBOSOMAL L10; PARTIAL.
 KW Ribosomal protein; Chloroplast; rRNA-binding.
 FT NON TER
 SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB0000 CRC64;
 Query Match 13.8%; Score 4; DB 10; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1e+06;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 S 1
 DB 3 S 3

RESULT 8
 ID P82182 PRELIMINARY; PRT; 6 AA.
 AC P82182
 DT 01-JUN-2000 (TrEMBLrel. 14, Created)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Chloroplast 50S ribosomal protein L10 gamma (Fragment).
 OS Spinacia oleracea (Spinach).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllales; Amaranthaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=cv. ALVARO; TISSUE=Leaf;
 RX MEDLINE=20435798; PubMed=10874046;
 RA Yamaguchi K., Subramanian A.R.;
 RT "The plastid ribosomal proteins. Identification of all the proteins in
 the 50 S subunit of an organelle ribosome (chloroplast).";
 RL J. Biol. Chem. 275:28466-28482 (2000).
 CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
 CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
 CC -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
 DR GO; GO:0009507; C:chloroplast; IEA.
 DR GO; GO:0019843; F:rRNA binding; IEA.
 DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
 DR InterPro; IPR002363; Ribosomal L10eub.
 DR PROSITE; PS01109; RIBOSOMAL L10; PARTIAL.
 KW Ribosomal protein; Chloroplast; rRNA-binding.
 FT NON TER
 SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB0000 CRC64;
 Query Match 13.8%; Score 4; DB 10; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1e+06;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 S 1
 DB 3 S 3

RESULT 9
 ID P83073 PRELIMINARY; PRT; 5 AA.
 AC P83073
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 8S kDa protein (Fragment).
 OS Bacillus cereus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1396;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=NCIMB 11796;
 RA Browne N., Dowds B.C.A.;
 RL Submitted (JUL-2001) to Swiss-Prot.
 FT NON TER
 SQ SEQUENCE 5 AA; 623 MW; 6B01AAA336F00000 CRC64;
 Query Match 6.9%; Score 2; DB 2; Length 5;
 Best Local Similarity 0.0%; Pred. No. 1e+06;
 Matches 0; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 6 L 6
 DB 1 M 1

RESULT 10

```

Q99007          PRELIMINARY;      PRT;      5 AA.
ID Q99007;
AC Q99007;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Alpha amylase (Fragment).
GN AMY1 GENE.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91329704; PubMed=1831055;
RA Jacobsen J.V., Close T.J.;
RT "Control of transient expression of chimeric genes by gibberellic
RT acid and abscisic acid in protoplasts prepared from mature barley
RT aleurone layers."
RL Plant Mol. Biol. 16:713-721(1991).
DR EMBL: X54643; CAA38455.1; -.
FT NON_TER 5
SQ SEQUENCE 5 AA; 600 MW; 61E3344DD6F00000 CRC64;

Query Match      6.9%; Score 2; DB 10; Length 5;
Best Local Similarity 0.0%; Pred. No. 1e+06;
Matches 0; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SS 2
Db ::
  2 AN 3

RESULT 11
P83533          PRELIMINARY;      PRT;      6 AA.
ID P83533;
AC P83533;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Unknown protein from 2D-page (Fragment).
OS Lactobacillus sanfranciscensis (Lactobacillus sanfrancisco).
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1625;
RN [1]
RP SEQUENCE.
RC STRAIN=DSM 20451;
RX PubMed=12112860;
RA Drews O., Weiss W., Reil G., Farlar H., Wait R., Georg A.;
RT "High pressure effects stepwise altered protein expression in
RT Lactobacillus sanfranciscensis."
RL Proteomics 2:765-774(2002).
CC -|- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED MW OF THIS UNKNOWN
CC PROTEIN IS: 15 KDA.
FT NON_TER 1
FT NON_TER 6
SQ SEQUENCE 6 AA; 590 MW; 6DDDD452D1AAC000 CRC64;

Query Match      6.9%; Score 2; DB 2; Length 6;
Best Local Similarity 0.0%; Pred. No. 1e+06;
Matches 0; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SS 2
Db 4 NA 5

```

Search completed: March 10, 2004, 15:15:31
 Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 15:12:32 ; Search time 20 Seconds
(without alignments)
28.857 Million cell updates/sec

Title: US-09-848-834A-6

Perfect score: 29

Sequence: 1 SSGPSL 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 316

Minimum DB seq length: 0

Maximum DB seq length: 6

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	62.1	4	2 PT0675	T-cell receptor be
2	17	58.6	5	2 JT0520	IG kappa chain V-I
3	17	58.6	5	2 PT0669	T-cell receptor be
4	14	48.3	4	2 PT0698	T-cell receptor be
5	14	48.3	5	2 PT0267	IG heavy chain CRD
6	14	48.3	5	2 PT0701	T-cell receptor be
7	14	48.3	5	2 PT0717	T-cell receptor be
8	14	48.3	6	2 PT0532	T-cell receptor be
9	14	48.3	6	2 PT0668	T-cell receptor be
10	14	48.3	6	2 PT0718	T-cell receptor be
11	14	48.3	6	2 PT0593	T-cell receptor be
12	13	44.8	6	2 A61049	halo-toxin - Pseud
13	13	44.8	6	2 S14159	parasporal crystal
14	13	44.8	6	2 JT0355	lipopeptide WS1279
15	12	41.4	5	2 PT0714	T-cell receptor be
16	12	41.4	6	2 PT0618	T-cell receptor be
17	12	41.4	6	2 PT0715	T-cell receptor be
18	11	37.9	3	3 PT0571	T-cell receptor be
19	11	37.9	4	2 PT0240	IG heavy chain CRD
20	11	37.9	4	2 PT0633	T-cell receptor be
21	11	37.9	4	2 PT0711	T-cell receptor be
22	11	37.9	4	2 PT0677	T-cell receptor be
23	11	37.9	4	2 PT0706	T-cell receptor be
24	11	37.9	4	2 PT0566	T-cell receptor be
25	11	37.9	5	2 C41225	copper resistance
26	11	37.9	5	2 A41225	copper resistance
27	11	37.9	5	2 S51077	alpha-amylase - ri
28	11	37.9	5	2 I50385	myosin light chain
29	11	37.9	5	2 PT0513	T-cell receptor be

ALIGNMENTS

RESULT 1

PT0675

T-cell receptor beta chain V-D-J region (140-14C) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C;Accession: PT0675

R;Reevey, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Reference number: PT0509; MUID:91277601; PMID:1711558

A;Accession: PT0675

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-4 <FEE>

A;Experimental source: day 18 fetal thymus, strain BALB/c

C;Keywords: T-cell receptor

Query Match 62.1%; Score 18; DB 2; Length 4;

Best Local Similarity 75.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSGP 4

Db :|||

1 ASGP 4

RESULT 2

JT0520

IG kappa chain V-III region (SD1) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 23-Oct-1992 #sequence_revision 23-Oct-1992 #text_change 16-Aug-1996

C;Accession: JT0520

R;Arker, R.; Conley, M.E.; Pollok, B.A.

J. Exp. Med. 169, 2109-2119, 1989

A;Title: Clonal diversity in the B cell repertoire of patients with X-linked agammaglobulinemia

A;Reference number: JT0511; MUID:89279157; PMID:2786547

A;Accession: JT0520

A;Molecule type: mRNA

A;Residues: 1-5 <ANK>

A;Note: the sequence shown here is one of eight productive V-D-J mu chain rearrangements

C;Note: a stop codon terminates the sequence in the V region

C;Keywords: heterotetramer; immunoglobulin

F;1-5/Domain: V kappa region <VRE>

Query Match

Best Local Similarity 58.6%; Score 17; DB 2; Length 5;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SGP 4

Db :|||

3 SGP 5

RESULT 3

PT0669
T-cell receptor beta chain V-D-J region (121-3BH) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0669
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0669
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-5 <FE2>
A:Experimental source: day 4 postnatal thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 56.6%; Score 17; DB 2; Length 5;

Best Local Similarity 75.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SSG 4

Db 2 SSG 5

RESULT 4

PT0698
T-cell receptor beta chain V-D-J region (126-1AH) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Sep-1997
C:Accession: PT0546; PT0698; PT0593
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0546
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-4 <FE2>
A:Experimental source: day 18 fetal thymus, strain BALB/c, 126-1AH
A:Accession: PT0698
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-4 <FE2>
A:Experimental source: newborn thymus, strain BALB/c, 135-1BP
A:Accession: PT0593
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-4 <FE3>
A:Experimental source: day 19 fetal thymus, strain BALB/c, clone 159-2C
C:Keywords: T-cell receptor

Query Match 48.3%; Score 14; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSG 3

Db 2 SSG 4

RESULT 5

PT0267
Ig heavy chain CDR3 region (clone 3-94A) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0267
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
A:Reference number: PT0222; MUID:91108337; PMID:1899102

A:Accession: PT0267
A:Molecule type: DNA
A:Residues: 1-5 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match 48.3%; Score 14; DB 2; Length 5;

Best Local Similarity 50.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SSG 4

Db 1 AGGP 4

RESULT 6

PT0701
T-cell receptor beta chain V-D-J region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0658; PT0701
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0658
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-5 <FE2>
A:Experimental source: day 4 postnatal thymus, strain BALB/c, 121-3E
A:Accession: PT0701
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-5 <FE2>
A:Experimental source: newborn thymus, strain BALB/c, 161-2E
C:Keywords: T-cell receptor

Query Match 48.3%; Score 14; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSG 3

Db 2 SSG 4

RESULT 7

PT0717
T-cell receptor beta chain V-D-J region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0717; PT0681
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0717
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-5 <FE2>
A:Experimental source: newborn thymus, strain BALB/c, 140-2B
A:Accession: PT0681
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-5 <FE2>
A:Experimental source: day 18 fetal thymus, strain BALB/c, 154-2C
C:Keywords: T-cell receptor

Query Match 48.3%; Score 14; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSG 3

Db 2 SSG 4
|||
Query Match 48.3%; Score 14; DB 2; Length 6;
Best Local Similarity 100.0%; Pred.No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 8
PT0532
T-cell receptor beta chain V-D-J region (100-4AJ) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0532
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0532
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FEE>
A:Experimental source: adult thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 48.3%; Score 14; DB 2; Length 6;
Best Local Similarity 100.0%; Pred.No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSG 3
|||
Db 2 SSG 4

RESULT 9
PT0668
T-cell receptor beta chain V-D-J region (121-3BB) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0668
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0668
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FEE>
A:Experimental source: day 4 postnatal thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 48.3%; Score 14; DB 2; Length 6;
Best Local Similarity 100.0%; Pred.No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSG 3
|||
Db 2 SSG 4

RESULT 10
PT0718
T-cell receptor beta chain V-D-J region (140-2D) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0718
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0718
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-6 <FEE>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 44.8%; Score 13; DB 2; Length 6;
Best Local Similarity 100.0%; Pred.No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 GP 4
|||
Db 4 GP 5

RESULT 11
PT0593
T-cell receptor beta chain V-D-J region (159-1F) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0593
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0593
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FEE>
A:Experimental source: day 19 fetal thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 48.3%; Score 14; DB 2; Length 6;
Best Local Similarity 100.0%; Pred.No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSG 3
|||
Db 3 SSG 5

RESULT 12
A61049
halo-toxin - Pseudomonas syringae pv. mori
C:Species: Pseudomonas syringae pv. mori
A:Note: host mulberry tree
C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 21-Jan-1997
C:Accession: A61049
R:Kajimoto, T.; Yokomizo, K.; Yahiro, K.; Umeda, T.; Shoji, S.; Kubota, Y.; Shibata, N.
Chem. Lett. 00, 679-680, 1989
A:Title: Structure of halo-toxin produced by phytopathogenic bacterium, Pseudomonas sy
A:Reference number: A61049
A:Accession: A61049
A:Molecule type: protein
A:Residues: 1-6 <KAJ>
A:Note: sequence confirmed by synthesis
C:Comment: This toxin is one of the etiological agents of halo bright diseases in mulbe
C:Keywords: toxin

Query Match 44.8%; Score 13; DB 2; Length 6;
Best Local Similarity 100.0%; Pred.No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 GP 4
|||
Db 4 GP 5

RESULT 13
S14159
parasporal crystal protein CryIC - Bacillus thuringiensis (fragment)
A:Alternate names: delta-endotoxin
C:Species: Bacillus thuringiensis
C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 18-Jul-1997
C:Accession: S14159
R:Convents, D.; Cherlet, M.; van Damme, J.; Lesters, I.; Lauwereys, M.
Eur. J. Biochem. 195, 631-635, 1991
A:Title: Two structural domains as a general fold of the toxic fragment of the Bacillu
A:Reference number: S14087; MUID:91153300; PMID:1847865

A:Accession: S14159
A:Molecule type: Protein
A:Residues: 1-6 <CON>

Query Match 44.8%; Score 13; DB 2; Length 6;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSGPS 5
|:|:
Db 2 STGNS 6

RESULT 14

JU0355
Lipopeptide WS1279 [validated] - Streptomyces willmorei
C:Species: Streptomyces willmorei
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: JU0355
R/Tsuda, Y.; Okada, Y.; Tanaka, M.; Shigematsu, N.; Hori, Y.; Goto, T.; Hashimoto, M.
Chem. Pharm. Bull. 39, 607-611, 1991
A>Title: Structure and synthesis of an immunoactive lipopeptide, WS1279, of microbial origin
A:Reference number: JU0355; PMID:91300586; PMID:2070441
A:Accession: JU0355
A:Molecule type: protein
A:Residues: 1-6 <TSU>
A>Note: the structure was confirmed by synthesis
C:Keywords: blocked amino end; lipoprotein
P:1/Binding site: sn-2,3-diacylglycerol (Cys) (covalent) #status experimental
P:1/Modified site: fatty acylated amino end (Cys) #status experimental

Query Match 44.8%; Score 13; DB 2; Length 6;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSGPS 5
|:|:
Db 2 NSGGS 6

RESULT 15

PT0714
T-cell receptor beta chain V-D-J region (165-31) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0714
R/Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A>Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; PMID:91277601; PMID:1711558
A:Accession: PT0714
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-5 <FEE>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 41.4%; Score 12; DB 2; Length 5;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSGP 4
|:|:
Db 1 ASSP 4

Search completed: March 10, 2004, 15:16:04
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 15:09:06 ; Search time 11 Seconds
(without alignments)
28.402 Million cell updates/sec

Title: US-09-848-834A-6

Perfect score: 29

Sequence: 1 SSGPSL 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 50

Minimum DB seq length: 0

Maximum DB seq length: 6

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13	44.8	5	1	UXA4 CHLTR
2	9	31.0	5	1	SUGA ACHDO
3	8	27.6	5	1	AL14 CARMA
4	8	27.6	5	1	BIOB_CITFR
5	8	27.6	5	1	PRCT_PEPAM
6	8	27.6	6	1	LOXI_LOCM1
7	8	27.6	6	1	WOPF_SARBU
8	8	27.6	6	1	VP19_HSVIK
9	7	24.1	3	1	THYL_PIG
10	7	24.1	4	1	DCML_PSECH
11	7	24.1	4	1	RM01_YEAST
12	7	24.1	4	1	TUFT_HUMAN
13	7	24.1	5	1	BPP7_BORIN
14	7	24.1	5	1	EIO3_LITRU
15	7	24.1	5	1	PAP2_PARMA
16	7	24.1	6	1	CIP1_MYTED
17	7	24.1	6	1	CIP2_MYTED
18	7	24.1	6	1	EIO1_LITRU
19	7	24.1	6	1	OM1_LIPDE
20	7	24.1	6	1	TRP1_PSEPU
21	6	20.7	3	1	GRW1_HUMAN
22	6	20.7	4	1	ACH1_ACHFU
23	6	20.7	4	1	BOS1_HUMAN
24	6	20.7	4	1	OCP1_OCTMI
25	6	20.7	4	1	OCP3_OCTMI
26	6	20.7	5	1	RE32_LITRU
27	6	20.7	5	1	TPIS_CANFA
28	6	20.7	5	1	UF01_MOUSE
29	6	20.7	6	1	FARP_MONEK
30	4	13.8	4	1	FAR3_HIRME
31	4	13.8	4	1	FLRF_HIRME
32	4	13.8	4	1	FLRN_ATEL
33	4	13.8	6	1	ACPH_RABIT

RESULT 1
ID UXA4 CHLTR STANDARD; PRT; 5 AA.
AC P38005;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Unknown protein from 2b-page from elementary body (Fragment).
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE.
RC STRAIN=L2/434/Bu;
RA Bini L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M., Comanducci M., Christianen G., Birkelund S., Vitreou E., Ratti G., Pallini V.;
RA Submitted (SEP-1994) to Swiss-Prot.
RL -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN PROTEIN IS: 4.5. ITS MW IS: 28 kDa.
CC Siena-2DPAGE; P38005; -.
DR NON-TER 5
FT SEQUENCE 5 AA; 474 MW; 75BAA865AA800000 CRC64;

Query Match 44.8%; Score 13; DB 1; Length 5;
Best Local Similarity 75.0%; Pred. No. 1.4e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 SGPS 5
DB 2 SGDS 5

RESULT 2
ID SUGA ACHDO STANDARD; PRT; 5 AA.
AC P19991;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Suboesophageal ganglion pentapeptide.
OS Acheta domestica (House cricket).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Gryllidae; Gryllinae;
OC Acheta.
RN [1]
RP SEQUENCE.
RA Wicker C., Wicker C.;
RT "Isolation and structure of a peptide isolated from the suboesophageal ganglion of Acheta domestica (orthoptera).";
RL Comp. Biochem. Physiol. 88C:185-187(1987).
CC -!- SUBCELLULAR LOCATION: Main peptide from the suboesophageal ganglia.
DR PIR: J50319; J50319.
SQ SEQUENCE 5 AA; 476 MW; 69D76DDDDDB00000 CRC64;

34 2 6.9 3 1 LUXE VIBFI
35 2 6.9 4 1 DCMS_PSECH
36 2 6.9 4 1 FAR4_HIRME
37 2 6.9 4 1 FMRP_MAGNI
38 2 6.9 4 1 FYRI_ATEL
39 2 6.9 5 1 B10A_CITFR
40 2 6.9 5 1 EIO4_LITRU
41 2 6.9 5 1 FARP_ARTTR
42 2 6.9 5 1 PSK_DAUCA
43 2 6.9 5 1 RE21_LITRU
44 2 6.9 5 1 RE31_LITRU
45 2 6.9 5 1 UC22_MAIZE

ALIGNMENTS

P24272 vibrio fisc
P19918 pseudomonas
P42563 hirudo medi
P01162 macrocallis
P58706 anthopleura
P13071 citrobacter
P82100 litorea rub
P41853 attoposthi
P58261 daucus caro
P82071 litorea rub
P82072 litorea rub
P80628 zea mays (m

Query Match 31.0%; Score 9; DB 1; Length 5;
 Best Local Similarity 25.0%; Pred. No. 1.4e+05;
 Matches 1; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSGP 4
 : :
 1 AAP 4

Db

RESULT 3
 AL14_CARMA STANDARD; PRT; 5 AA.
 AC P81817;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 14;
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RA Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
 CC -!- SIMILARITY: Belongs to the allatostatin family.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD RES 5
 FT AMIDATION (POTENTIAL).
 SQ SEQUENCE 5 AA; 586 MW; 672879D5AB300000 CRC64;

Query Match 27.6%; Score 8; DB 1; Length 5;
 Best Local Similarity 86.%; Pred. No. 1.4e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSG 3
 : :
 2 SFG 4

Db

RESULT 4
 BIOB CITFR STANDARD; PRT; 5 AA.
 ID BIOB CITFR
 AC P12997;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Biotin synthase (EC 2.8.1.6) (Biotin synthetase) (Fragment).
 GN BIOB.
 OS Citrobacter freundii.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Citrobacter.
 OX NCBI_TaxID=546;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9906280; PubMed=2971595;
 RA Shivan D., Campbell A.;
 RT "Transcriptional regulation and gene arrangement of Escherichia coli,
 RT Citrobacter freundii and Salmonella typhimurium biotin operons";
 RL Gene 67:203-211(1988).
 CC -!- CATALYTIC ACTIVITY: Dethiobiotin + sulfur = biotin.
 CC -!- PATHWAY: Biotin biosynthesis; last step.
 CC -!- SIMILARITY: Belongs to the biotin and lipoic acid synthetases
 family.
 CC
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M21922; ; NOT_ANNOTATED_CDS.
 DR PIR; I40698; I40698.
 KW Biotin biosynthesis; Iron-sulfur; Transferase.
 FT NON_TER 5
 SQ SEQUENCE 5 AA; 532 MW; 75ASB1EDD6F00000 CRC64;

Query Match 27.6%; Score 8; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.4e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SS 2
 : :
 4 SS 5

Db

RESULT 5
 PRCT PERAM STANDARD; PRT; 5 AA.
 ID PRCT PERAM
 AC P01373; 1996 (Rel. 01, Created)
 DT 21-JUL-1996 (Rel. 01, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Proctolin.
 OS Periplaneta americana (American cockroach).
 OS Limulus polyphemus (Atlantic horseshoe crab). and
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattodea;
 OC Blattidae; Periplaneta.
 OX NCBI_TaxID=6978, 6850, 6759;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=P.americana;
 RX MEDLINE=76074708; PubMed=576;
 RA Starratt A.N., Brown B.E.;
 RT "Structure of the pentapeptide proctolin, a proposed neurotransmitter
 RT in insects.";
 RL Life Sci. 17:1253-1256(1975).
 RN [2]
 RP BIOLOGICAL SOURCE.
 RC SPECIES=P.americana;
 RX MEDLINE=81225865; PubMed=6113690;
 RA O'Shea M., Adams M.E.;
 RT "Pentapeptide (proctolin) associated with an identified neuron.";
 RL Science 213:567-569(1981).
 RN [3]
 RP SEQUENCE.
 RC SPECIES=L.polyphemus;
 RX MEDLINE=90287800; PubMed=2356151;
 RA Groome J.R., Tillinghast E.K., Townley M.A., Vetrovs A.,
 RA Watson W.H. III, Hunt D.F., Griffin P.R., Alexander J.E.,
 RA Shabanowitz J.;
 RT "Identification of proctolin in the central nervous system of the
 RT horseshoe crab, Limulus polyphemus.";
 RL Peptides 11:205-211(1990).
 RN [4]
 RP SEQUENCE.
 RC SPECIES=C.maenas;
 RX MEDLINE=86232789; PubMed=2872661;
 RA Stangier J., Dircksen H., Keller R.;
 RT "Identification and immunocytochemical localization of proctolin in
 RT pericardial organs of the shore crab, Carcinus maenas.";
 RL Peptides 7:67-72(1986).
 CC -!- FUNCTION: Stimulates cardiac output and hindgut motility,
 CC modulates visceral and skeletal muscle in many arthropods.
 CC -!- TISSUE SPECIFICITY: Found in the lateral white neurons and in

ID THYL_PIG STANDARD; PRT; 3 AA.
 AC P01151;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Thyroliberin (Thyrotropin releasing hormone) (TRH) (Protirelin).
 OS Sus scrofa (Pig).
 OS Ovis aries (Sheep).
 OS Bombina orientalis (Oriental fire-bellied toad), and
 OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9623, 9940, 8346, 8316;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=Pig; TISSUE=Hypothalamus;
 RX MEDLINE=70136150; PubMed=4984938;
 RA Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;
 RT "Structure of porcine thyrotropin releasing hormone.";
 RL Biochemistry 9:1103-1106(1970).
 RN [2]
 RP SYNTHESIS.
 RC SPECIES=Pig;
 RX MEDLINE=70039904; PubMed=4982117;
 RA Boler J., Enzmann F., Folkers K., Bowers C.Y., Schally A.V.;
 RT "The identity of chemical and hormonal properties of the thyrotropin
 releasing hormone and pyroglutamyl-histidyl-proline amide.";
 RL Biochem. Biophys. Res. Commun. 37:705-710(1969).
 RN [3]
 RP SEQUENCE.
 RC SPECIES=Sheep; TISSUE=Hypothalamus;
 RA Desiderio D.M. Jr., Burgess R., Dunn T.F., Vale W., Guillemin R.,
 Ward D.N.;
 RT "The elucidation of the primary structure of the hypothalamic thyroid
 stimulating hormone releasing factor of ovine origin by means of mass
 spectrometry.";
 RL Org. Mass Spectrom. 5:221-228(1971).
 RN [4]
 RP SYNTHESIS.
 RC SPECIES=Sheep;
 RX MEDLINE=70163386; PubMed=4985794;
 RA Burgess R., Dunn T.F., Desiderio D.M., Ward D.N., Vale W.,
 Guillemin R.;
 RT "Characterization of ovine hypothalamic hypophysiotropic
 TSH-releasing factor.";
 RL Nature 226:321-325(1970).
 RN [5]
 RP SEQUENCE.
 RC SPECIES=B. orientalis; TISSUE=Skin;
 RX MEDLINE=76138399; PubMed=815011;
 RA Yasuhara T., Nakajima T.;
 RT "Letter: Occurrence of Pyr-His-Pro-NH2 in the frog skin.";
 RL Chem. Pharm. Bull. 23:3301-3303(1975).
 RN [6]
 RP SEQUENCE.
 RC SPECIES=N. viridescens;
 RX MEDLINE=75035605; PubMed=4214528;
 RA Grimm-Jorgensen Y., McKelvy J.E.;
 RT "Biosynthesis of thyrotropin releasing factor by newt (Triturus
 viridescens) brain in vitro. Isolation and characterization of
 thyrotropin releasing factor.";
 RL J. Neurochem. 23:471-478(1974).
 CC -1- FUNCTION: TRH functions as a regulator of the biosynthesis of TSH
 in the anterior pituitary gland and as a neurotransmitter/
 neuroendulator in the central and peripheral nervous systems.
 DR PIR; A90919; RHDT0.
 DR PIR; A92971; A92971.
 DR PIR; A93750; RHSH.
 KW Amidation; Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 3 3 AMIDATION.
 FT MOD_RES 3 3
 SQ SEQUENCE 3 AA; 380 MW; 776186B000000000 CRC64;

Query Match 24.1%; Score 7; DB 1; Length 3;
 Best Local Similarity 100.0%; Pred. No. 1.4e+05; Indels 0; Gaps 0;
 Matches 1; Conservative 0; Mismatches 0;
 Qy 4 P 4
 Db 3 P 3
 RESULT 10
 DCML_PSECH STANDARD; PRT; 4 AA.
 ID DCML_PSECH
 AC P19916;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) (CO
 dehydrogenase subunit L) (CO-DH L) (Fragment).
 GN CMTL.
 OS Pseudomonas carboxydohydrogena.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae.
 OX NCBI_TaxID=290;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=30055678; PubMed=2818128;
 RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
 RT "Homology and distribution of CO dehydrogenase structural genes in
 carboxydohydrogenic bacteria.";
 RL Arch. Microbiol. 152:335-341(1989).
 CC -1- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
 dioxide.
 CC -1- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
 acceptor.
 CC -1- COFACTOR: Molybdenum (molybdopterin).
 CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
 SMALL.
 DR PIR; F01440; F01440.
 KW Oxidoreductase; Molybdenum.
 FT NON_TER 4
 SQ SEQUENCE 4 AA; 441 MW; 77618676F000000000 CRC64;
 Query Match 24.1%; Score 7; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.4e+05; Indels 0; Gaps 0;
 Matches 1; Conservative 0; Mismatches 0;
 Qy 4 P 4
 Db 4 P 4
 RESULT 11
 RM01_YEAST STANDARD; PRT; 4 AA.
 ID RM01_YEAST
 AC P36515;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Mitochondrial 60S ribosomal protein L1 (Yml1) (Fragment).
 GN MRPL1.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=91285105; PubMed=2060626;
 RA Grohmann L., Graack H.-R., Kruft V., Choli T., Goldschmidt-Reisin S.,
 Kitakawa M.;
 RT "Extended N-terminal sequencing of proteins of the large ribosomal
 subunit from yeast mitochondria.";
 RL FEBS Lett. 284:51-56(1991).
 DR PIR; S17255; S17255.

```

DR SGD; L0002681; MRPL1.
KW Ribosomal protein; Mitochondrion.
FT NON TER 4
SQ SEQUENCE 4 AA; 402 MW; 7771B2D5D0000000 CRC64;

Query Match 24.1%; Score 7; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 P 4
Db 4 P 4

RESULT 12
TUFT HUMAN
ID TUFT_HUMAN STANDARD; PRT; 4 AA.
AC P01658;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phagocytosis-stimulating peptide (Tuftsin).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=72187087; PubMed=4112769;
RA Nishioka K., Constantopoulos A., Satoh P.S., Najjar V.A.;
RT "The characteristics, isolation and synthesis of the phagocytosis
RT stimulating peptide tuftsin."
RL Biochem. Biophys. Res. Commun. 47:172-179(1972).
RN [2]
RP IMMUNOGLOBULIN CLASS.
RX MEDLINE=68091045; PubMed=4169272;
RA Fidalgo B.V., Najjar V.A.;
RT "The physiological role of the lymphoid system. VI. The stimulatory
RT effect of leucophilic gamma globulin (leucokinin) on the phagocytic
RT activity of human polymorphonuclear leucocyte."
RL Biochemistry 6:3386-3392(1967).
CC -1- MISCELLANEOUS: An IgG (called leucokinin) binds reversibly to the
CC cell membrane of neutrophils in the blood. Leucokininase on the
CC membrane releases the active peptide tuftsin from the gamma chain.
CC Tuftsin is essential for maximum stimulation of the phagocytic
CC activity of neutrophils.
DR PIR; A02147; A02147.
DR MIM; 191150; -.
DR GO; GO:0003823; P:antigen binding; NAS.
DR GO; GO:0006909; P:phagocytosis; NAS.
SQ SEQUENCE 4 AA; 501 MW; 74176321C0000000 CRC64;

Query Match 24.1%; Score 7; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 P 4
Db 3 P 3

RESULT 13
BPP7_BOTIN
ID BPP7_BOTIN STANDARD; PRT; 5 AA.
AC P10425;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bradykinin-potentiating peptide S5,2 (5A) (Angiotensin-converting
DE enzyme inhibitor).
OS Botnops insularis (Island jararaca) (Queimada jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroides;

DR SGD; L0002681; MRPL1.
KW Ribosomal protein; Mitochondrion.
FT NON TER 4
SQ SEQUENCE 4 AA; 402 MW; 7771B2D5D0000000 CRC64;

Query Match 24.1%; Score 7; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 P 4
Db 4 P 4

RESULT 14
E103_LITRU
ID E103_LITRU STANDARD; PRT; 5 AA.
AC F82039;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Electrin 3.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RX TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
RT Litoria electrica. Comparison with the skin peptides from Litoria
RT rubella."
RL Aust. J. Chem. 52:639-645(1999).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Skin.
KW Amphibian defense peptide; Amidation.
FT MOD RES 5 5
SQ SEQUENCE 5 AA; 630 MW; 668761F2C9A00000 CRC64;

Query Match 24.1%; Score 7; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 P 4
Db 4 P 4

RESULT 15
PAP2_PARMA
ID PAP2_PARMA STANDARD; PRT; 5 AA.
AC P81864;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pardaxin II (P11) (Fragment).
OS Pardachirus marmoratus (Red sea mores sole).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Soleidae; Soleidae; Pardachirus.
 OX NCBI_TaxID=31087;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Skin secretion;
 RX MEDLINE=87057369; PubMed=3782138;
 RA Lazarovici P., Primor N., Loew L.M.;
 RT "Purification and pore-forming activity of two hydrophobic
 RT polypeptides from the secretion of the Red sea moose sole (Pardachirus
 RT marmoratus).";
 RL J. Biol. Chem. 261:16704-16713(1986).
 CC -!- FUNCTION: Exhibits unusual shark repellent and surfactant
 CC properties. Forms voltage-dependent, ion-permeable channels
 CC in membranes. At high concentration causes cell membrane lysis.
 CC -!- SUBUNIT: Monomer. In aqueous solution exists as a tetramer.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the pardaxin family.
 KW Toxin.
 FT NON_TER 5 5
 SQ SEQUENCE 5 AA; 614 MW; 7769C9C9C8100000 CRC64;

 Query Match 24.1%; Score 7; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.4e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 4 P 4
 Db 5 P 5

Search completed: March 10, 2004, 15:14:40
 Job time : 12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 15:05:36 ; Search time 23 Seconds
(without alignments)
8.978 Million cell updates/sec

Title: US-09-848-834A-5

Sequence: 1 GPLS 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 11020

Minimum DB seq length: 0
Maximum DB seq length: 4

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	85.7	4	1	US-08-206-789-2
2	16	76.2	4	1	US-07-932-200-8
3	16	76.2	4	2	US-08-387-749-8
4	16	76.2	4	5	PCT-US93-08231-8
5	15	71.4	4	1	US-08-206-789-1
6	15	71.4	4	1	US-08-213-124-18
7	15	71.4	4	2	US-08-685-589A-88
8	14	66.7	4	1	US-08-206-789-5
9	14	66.7	4	1	US-08-213-897A-3
10	14	66.7	4	1	US-08-329-820-77
11	14	66.7	4	1	US-08-329-820-83
12	14	66.7	4	2	US-08-685-589A-35
13	14	66.7	4	3	US-09-032-308A-14
14	14	66.7	4	3	US-08-836-480-1
15	14	66.7	4	3	US-09-562-897-1
16	14	66.7	4	4	US-09-584-001C-14
17	13	61.9	3	1	US-07-816-679A-9
18	13	61.9	3	1	US-08-365-759-1
19	13	61.9	3	1	US-08-213-897A-11
20	13	61.9	3	1	US-08-213-897A-12
21	13	61.9	3	2	US-08-476-044-8
22	13	61.9	3	4	US-09-521-365A-1
23	13	61.9	3	4	US-08-695-692B-1
24	13	61.9	3	4	US-09-755-630B-277
25	13	61.9	3	4	US-09-755-274-16
26	13	61.9	3	5	PCT-US92-11270-9
27	13	61.9	4	1	US-07-963-324-1

28 13 61.9 4 1 US-07-934-553-5 Sequence 5, Appli
29 13 61.9 4 1 US-07-923-692C-1 Sequence 1, Appli
30 13 61.9 4 1 US-07-820-688A-5 Sequence 5, Appli
31 13 61.9 4 1 US-07-932-200-6 Sequence 6, Appli
32 13 61.9 4 1 US-07-932-200-7 Sequence 9, Appli
33 13 61.9 4 1 US-07-932-200-9 Sequence 1, Appli
34 13 61.9 4 1 US-07-776-257-1 Sequence 2, Appli
35 13 61.9 4 1 US-07-776-257-2 Sequence 2, Appli
36 13 61.9 4 1 US-08-127-351-50 Sequence 50, Appli
37 13 61.9 4 1 US-07-618-946B-1 Sequence 27, Appli
38 13 61.9 4 1 US-08-079-812-27 Sequence 1, Appli
39 13 61.9 4 1 US-08-022-381A-1 Sequence 2, Appli
40 13 61.9 4 1 US-08-067-387-2 Sequence 6, Appli
41 13 61.9 4 1 US-08-432-617-6 Sequence 6, Appli
42 13 61.9 4 1 US-08-166-316-6 Sequence 17, Appli
43 13 61.9 4 1 US-08-243-082-17 Sequence 9, Appli
44 13 61.9 4 1 US-08-095-162-9 Sequence 5, Appli
45 13 61.9 4 1 US-08-219-156-5

ALIGNMENTS

RESULT 1
US-08-206-789-2
; Sequence 2, Application US/08206789
; Patent No. 5580854
; GENERAL INFORMATION:
; APPLICANT: Orłowski, Marian
; APPLICANT: Cardozo, Christopher
; APPLICANT: Vinitsky, Alexander
; TITLE OF INVENTION: SUBSTRATE-RELATED
; TITLE OF INVENTION: PEPTIDYL-ALDEHYDE
; TITLE OF INVENTION: INHIBITORS OF THE PROTEOLYTIC ACTIVITY OF THE
; TITLE OF INVENTION: MULTICATALYTIC PROTEINASE COMPLEX
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRUMBAUGH, GRAVES, DONOHUE &
; ADDRESSEE: RAYMOND
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112-2500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/206,789
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Seide, Rochelle K.
; REGISTRATION NUMBER: 32,300
; REFERENCE/DOCKET NUMBER: A29525 - 165/25989
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2626
; TELEFAX: 212-765-2519
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..4
; US-08-206-789-2

Query Match 85.7% ; Score 18; DB 1; Length 4;
Best Local Similarity 75.0% ; Pred. No. 3e+05;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPDL 4
|||
Db 1 GPAL 4

RESULT 2
US-07-932-200-8
; Sequence 8, Application US/07932200
; Patent No. 5366862
; GENERAL INFORMATION:
; APPLICANT: VENTON, DUANE L.
; APPLICANT: HOPFINGER, ANTON J.
; APPLICANT: LE BRETON, GUY
; TITLE OF INVENTION: METHOD FOR GENERATING AND SCREENING
; TITLE OF INVENTION: USEFUL PEPTIDES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 SEVENTH STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/932,200
; FILING DATE: 21-AUG-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, IVER P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: VENTON-1B
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-932-200-8

Query Match 76.2%; Score 16; DB 1; Length 4;
Best Local Similarity 75.0%; Pred. No. 3e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

Qy 1 GPDL 4
|||
Db 1 GPRL 4

RESULT 3
US-08-387-749-8
; Sequence 8, Application US/08387749
; Patent No. 581460
; GENERAL INFORMATION:
; APPLICANT: VENTON, DUANE L.
; APPLICANT: HOPFINGER, ANTON J.
; APPLICANT: LE BRETON, GUY
; TITLE OF INVENTION: METHOD FOR GENERATING AND SCREENING
; TITLE OF INVENTION: USEFUL PEPTIDES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK

STREET: 419 SEVENTH STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/387,749
; FILING DATE: 21-Feb-1995
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08231
; FILING DATE: 09-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/932,200
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA: US 07/813,315
; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/00891
; FILING DATE: 14-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/480,865
; FILING DATE: 14-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, IVER P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: VENTON-1C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-387-749-8

Query Match 76.2%; Score 16; DB 2; Length 4;
Best Local Similarity 75.0%; Pred. No. 3e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GPDL 4
|||
Db 1 GPRL 4

RESULT 4
PCT-US93-08231-8
; Sequence 8, Application PC/TUS9308231
; GENERAL INFORMATION:
; APPLICANT: VENTON, DUANE L.
; APPLICANT: HOPFINGER, ANTON J.
; APPLICANT: LE BRETON, GUY
; TITLE OF INVENTION: METHOD FOR GENERATING AND SCREENING
; TITLE OF INVENTION: USEFUL PEPTIDES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 SEVENTH STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US93/08231
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: COOPER, IVER P.
;; REGISTRATION NUMBER: 28,005
;; REFERENCE/DOCKET NUMBER: VENTONIC.PCT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-628-5197
;; TELEFAX: 202-737-3528
;; TELEX: 248633
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
PCT-US93-08231-8

Query Match 76.2%; Score 16; DB 5; Length 4;
Best Local Similarity 75.0%; Pred. No. 3e+05; 1; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 GPGL 4
Db 1 GPRL 4

RESULT 5
US-08-206-789-1
; Sequence 1, Application US/08206789
; Patent No. 5580854
; GENERAL INFORMATION:
; APPLICANT: Orłowski, Marian
; APPLICANT: Cardozo, Christopher
; APPLICANT: Vinitzky, Alexander
; TITLE OF INVENTION: SUBSTRATE-RELATED
; TITLE OF INVENTION: PEPTIDYL-ALDEHYDE
; TITLE OF INVENTION: INHIBITORS OF THE PROTEOLYTIC ACTIVITY OF THE
; TITLE OF INVENTION: MULTICATALYTIC PROTEINASE COMPLEX
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRUMBAUGH, GRAVES, DONOHUE &
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112-2500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/206.789
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Seide, Rochelle K.
; REGISTRATION NUMBER: 32,300
; REFERENCE/DOCKET NUMBER: A29525 - 165/25989
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2626
; TELEFAX: 212-765-2519
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

;; LENGTH: 4 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FEATURE:
;; NAME/KEY: Peptide
;; LOCATION: 1..4
US-08-206-789-1

Query Match 71.4%; Score 15; DB 1; Length 4;
Best Local Similarity 75.0%; Pred. No. 3e+05; 1; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 GPGL 4
Db 1 GPFL 4

RESULT 6
US-08-213-124-18
; Sequence 18, Application US/08213124
; Patent No. 5693325
; GENERAL INFORMATION:
; APPLICANT: Kahn, Michael
; TITLE OF INVENTION: PEPTIDE VACCINES AND METHODS RELATING
; TITLE OF INVENTION: THERETO
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/213.124
; FILING DATE: 15-MAR-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Hermanns, Karl R.
; REGISTRATION NUMBER: 33,507
; REFERENCE/DOCKET NUMBER: 670063.411
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836 SEEDANDBERRY
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-213-124-18

Query Match 71.4%; Score 15; DB 1; Length 4;
Best Local Similarity 50.0%; Pred. No. 3e+05; 1; Indels 0; Gaps 0;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GPGL 4
Db 1 GPGL 4

RESULT 7
US-08-685-589A-88
; Sequence 88, Application US/08685589A
; Patent No. 5916872
; GENERAL INFORMATION:

APPLICANT: Chang, Conway
APPLICANT: Gu, Leo
APPLICANT: Chen, Jie
TITLE OF INVENTION: CYCLIC PEPTIDES HAVING BROAD
TITLE OF INVENTION: SPECTRUM ANTIMICROBIAL ACTIVITY
NUMBER OF SEQUENCES: 222
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 24-JUL-1996
APPLICATION NUMBER: US/08/685,589A
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 8067-026-999
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: No. 516872 Relevant
MOLECULE TYPE: peptide
US-08-685-589A-88

Query Match 71.4%; Score 15; DB 2; Length 4;
Best Local Similarity 50.0%; Pred. No. 3e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GPSL 4
Db 1 GPQI 4

RESULT 8
US-08-206-789-5
Sequence 5, Application US/08206789
Patent No. 5580854
GENERAL INFORMATION:
APPLICANT: Orłowski, Marian
APPLICANT: Cardozo, Christopher
APPLICANT: Vinitzky, Alexander
TITLE OF INVENTION: SUBSTRATE-RELATED
TITLE OF INVENTION: PEPTIDYL-ALDEHYDE
TITLE OF INVENTION: INHIBITORS OF THE PROTEOLYTIC ACTIVITY OF THE
TITLE OF INVENTION: MULTICATALYTIC PROTEINASE COMPLEX
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRUMBAUGH, GRAVES, DONOHUE &
ADDRESSEE: RAYMOND
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112-2500
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/206,789
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Seide, Rochelle K.
REGISTRATION NUMBER: 32,300
REFERENCE/DOCKET NUMBER: A29525 - 165/25989
TELEPHONE: 212-408-2626
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..4
US-08-206-789-5

Query Match 66.7%; Score 14; DB 1; Length 4;
Best Local Similarity 66.7%; Pred. No. 3e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GPS 3
Db 1 GPA 3

RESULT 9
US-08-213-897A-3
Sequence 3, Application US/08213897A
Patent No. 5618790
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Protease Mediated Drug Delivery System
NUMBER OF SEQUENCES: 18
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/213,897A
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/593,867
FILING DATE: 05-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/833,183
FILING DATE: 10-FEB-1992
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-213-897A-3

Query Match 66.7%; Score 14; DB 1; Length 4;
Best Local Similarity 66.7%; Pred. No. 3e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GPS 3
Db 2 GPA 4

```
RESULT 10
US-08-329-820-77
; Sequence 77, Application US/08329820
; Patent No. 5747457
; GENERAL INFORMATION:
; APPLICANT: JONCZYK, ALFRED
; APPLICANT: FELDING-HABERMANN, BRUNHILDE
; APPLICANT: DIEPENBACH, BEATE
; APPLICANT: RIPPWANN, FRIEDRICH
; TITLE OF INVENTION: Linear Adhesion Inhibitors
; NUMBER OF SEQUENCES: 228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/329,820
; FILING DATE: 27-OCT-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4336758.5
; FILING DATE: 28-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hamlet-King, Diana
; REGISTRATION NUMBER: 33,302
; REFERENCE/DOCKET NUMBER: MERCK 1635
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: linear
US-08-329-820-77

Query Match 66.7%; Score 14; DB 1; Length 4;
Best Local Similarity 66.7%; Pred. No. 3e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPS 3
Db 2 GPA 4

RESULT 11
US-08-329-820-83
; Sequence 83, Application US/08329820
; Patent No. 5747457
; GENERAL INFORMATION:
; APPLICANT: JONCZYK, ALFRED
; APPLICANT: FELDING-HABERMANN, BRUNHILDE
; APPLICANT: DIEPENBACH, BEATE
; APPLICANT: RIPPWANN, FRIEDRICH
; TITLE OF INVENTION: Linear Adhesion Inhibitors
; NUMBER OF SEQUENCES: 228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/329,820
; FILING DATE: 27-OCT-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4336758.5
; FILING DATE: 28-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hamlet-King, Diana
; REGISTRATION NUMBER: 33,302
; REFERENCE/DOCKET NUMBER: MERCK 1635
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: linear
US-08-329-820-77

Query Match 66.7%; Score 14; DB 1; Length 4;
Best Local Similarity 66.7%; Pred. No. 3e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPS 3
Db 2 GPA 4

RESULT 12
US-08-685-589A-35
; Sequence 35, Application US/08685589A
; Patent No. 5916872
; GENERAL INFORMATION:
; APPLICANT: Chang, Conway
; APPLICANT: Gu, Leo
; APPLICANT: Chen, Jie
; TITLE OF INVENTION: CYCLIC PEPTIDES HAVING BROAD
; TITLE OF INVENTION: SPECTRUM ANTIMICROBIAL ACTIVITY
; NUMBER OF SEQUENCES: 222
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,589A
; FILING DATE: 24-JUL-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 8067-026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; APPLICATION NUMBER: DE 4336758.5
; FILING DATE: 28-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hamlet-King, Diana
; REGISTRATION NUMBER: 33,302
; REFERENCE/DOCKET NUMBER: MERCK 1635
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: linear
US-08-329-820-83

Query Match 66.7%; Score 14; DB 1; Length 4;
Best Local Similarity 66.7%; Pred. No. 3e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPS 3
Db 1 GPA 3
```

; TOPOLOGY: No. 5916872 Relevant
 ; MOLECULE TYPE: peptide
 ; FEATURE:
 ; NAME/KEY: Peptide
 ; LOCATION: 1..4
 ; OTHER INFORMATION: /product= "Beta-turn"
 ; US-08-685-589A-35

Query Match 66.7%; Score 14; DB 2; Length 4;
 Best Local Similarity 50.0%; Pred. No. 3e+05;
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GPSL 4
 ||:
 Db 1 GPRI 4

RESULT 13
 US-09-039-308A-14
 ; Sequence 14, Application US/09039308A
 ; Patent No. 6069129
 ; GENERAL INFORMATION:
 ; APPLICANT: Sandberg, Lawrence; Roos, Phillip;
 ; APPLICANT: Mitts, Thomas
 ; TITLE OF INVENTION: ELASTIN DERIVED COMPOSITION
 ; TITLE OF INVENTION: AND METHOD OF
 ; TITLE OF INVENTION: USING SAME
 ; NUMBER OF SEQUENCES: 41
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: REED SMITH SHAW & MCCRAY, LLP
 ; STREET: PO Box 488
 ; CITY: Pittsburgh
 ; STATE: Pennsylvania
 ; COUNTRY: USA
 ; ZIP: 15230

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb
 ; COMPUTER: Compaq
 ; OPERATING SYSTEM: Microsoft Windows 95
 ; SOFTWARE: Word 6.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/039,308A
 ; FILING DATE: March 13, 1998
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Miller, Raymond A.
 ; REGISTRATION NUMBER: 42,891
 ; REFERENCE/DOCKET NUMBER: 97-489
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (412) 288-4192
 ; TELEFAX: (412) 288-3300

; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-09-039-308A-14

Query Match 66.7%; Score 14; DB 3; Length 4;
 Best Local Similarity 66.7%; Pred. No. 3e+05;
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPS 3
 ||:
 Db 2 GPA 4

RESULT 14
 US-08-836-480-1
 ; Sequence 1, Application US/08836480
 ; Patent No. 6103697

; GENERAL INFORMATION:
 ; APPLICANT: Bergstrand, Hakan
 ; APPLICANT: Erickson, Tomas
 ; APPLICANT: Karabelas, Kostas
 ; APPLICANT: Lindvall, Magnus
 ; APPLICANT: Sarnstrand, Bengt
 ; TITLE OF INVENTION: NEW PEPTIDES WITH IMMUNOMODULATORY EFFECTS
 ; NUMBER OF SEQUENCES: 39
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson, P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: US
 ; ZIP: 02110-2804

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/836,480
 ; FILING DATE: 14-APR-1997
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: SE 95\403256-8
 ; FILING DATE: 14-OCT-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fraser, Janis K.
 ; REGISTRATION NUMBER: 34,819
 ; REFERENCE/DOCKET NUMBER: 06275/036001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-542-5070
 ; TELEFAX: 617-542-8906

; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-836-480-1

Query Match 66.7%; Score 14; DB 3; Length 4;
 Best Local Similarity 50.0%; Pred. No. 3e+05;
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GPSL 4
 ||:
 Db 1 GPCI 4

RESULT 15
 US-09-562-897-1
 ; Sequence 1, Application US/09562897
 ; Patent No. 6228374

; GENERAL INFORMATION:
 ; APPLICANT: Bergstrand, Hakan
 ; Erickson, Tomas
 ; Karabelas, Kostas
 ; Lindvall, Magnus
 ; Sarnstrand, Bengt

; TITLE OF INVENTION: NEW PEPTIDES WITH IMMUNOMODULATORY EFFECTS
 ; NUMBER OF SEQUENCES: 39
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson, P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: US
 ; ZIP: 02110-2804

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/562,897
FILING DATE: 01-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,480
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 06275/036001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-562-897-1

Query Match 66.7%; Score 14; DB 3; Length 4;
Best Local Similarity 50.0%; Pred. No. 3e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GPSL 4
|||
Db 1 GPCI 4

Search completed: March 10, 2004, 15:09:00
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 15:04:45 ; Search time 21 Seconds
(without alignments)
18.322 Million cell updates/sec

Title: US-09-848-834A-5

Perfect score: 21

Sequence: 1 GPGL 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 86

Minimum DB seq length: 0

Maximum DB seq length: 4

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13	61.9	4	2 PT0675	T-cell receptor be
2	11	52.4	4	2 PT0240	Ig heavy chain CRD
3	9	42.9	4	2 I54357	schwannomin - mous
4	8	38.1	3	3 I78890	tyrosine protein k
5	8	38.1	4	2 S3508	starvation-induced
6	8	38.1	4	2 PT0645	T-cell receptor be
7	7	33.3	3	3 RHTDFO	thyroliberin - Bom
8	7	33.3	3	3 RHPGT	thyroliberin - pig
9	7	33.3	3	3 RUSHT	thyroliberin - she
10	7	33.3	3	3 A92971	thyroliberin - eas
11	7	33.3	3	3 A33802	thyrotropin-releas
12	7	33.3	3	3 A43391	TRH-like tripeptid
13	7	33.3	4	2 A02147	phagocytosis-stimu
14	7	33.3	4	2 A32039	tyrosine-melanocyt
15	7	33.3	4	2 PL0140	carbon-monoxide de
16	7	33.3	4	2 S17255	ribosomal protein
17	7	33.3	4	2 A34626	RPCH-related neuro
18	7	33.3	4	2 I51049	metallothionein-A
19	6	28.6	3	3 GKHU	growth-modulating
20	6	28.6	3	3 A60898	bursin - chicken
21	6	28.6	3	3 A23751	spinal cord peptid
22	6	28.6	3	3 B23751	spinal cord peptid
23	6	28.6	3	3 PT0636	T-cell receptor be
24	6	28.6	3	3 PT0578	T-cell receptor be
25	6	28.6	3	3 PT0571	T-cell receptor be
26	6	28.6	4	1 ECXAA	antho-RFamide neur
27	6	28.6	4	2 T30569	hypothetical prote
28	6	28.6	4	2 I38888	COI intron 16 prot
29	6	28.6	4	2 A25844	autho-RF amide neu

30 6 28.6 4 2 A32480
31 6 28.6 4 2 S39390
32 6 28.6 4 2 PT0271
33 6 28.6 4 2 S43959
34 6 28.6 4 2 A53284
35 6 28.6 4 2 B53284
36 6 28.6 4 2 PT0633
37 6 28.6 4 2 PT0711
38 6 28.6 4 2 PT0698
39 6 28.6 4 2 PT0677
40 6 28.6 4 2 PT0551
41 6 28.6 4 2 PT0697
42 6 28.6 4 2 PT0706
43 6 28.6 4 2 PT0721
44 6 28.6 4 2 PT0586
45 6 28.6 4 2 S47552

achatin-I - giant
myosin-light-chain
Ig heavy chain CRD
Ig mu chain V regi
T-cell receptor be
T-cell receptor be
T-cell receptor be
T-cell receptor be
T-cell receptor be
T-cell receptor be
T-cell receptor be
T-cell receptor be
T-cell receptor be
ubiquitin - rat

ALIGNMENTS

RESULT 1

PT0675 T-cell receptor beta chain V-D-J region (140-14C) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C;Accession: PT0675

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Reference number: PT0509; MUID:91277601; PMID:1711558

A;Accession: PT0675

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-4 <FE>

A;Experimental source: day 18 fetal thymus, strain BALB/c

C;Keywords: T-cell receptor

Query Match 61.9%; Score 13; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GP 2

Db 3 GP 4

RESULT 2

PT0240

Ig heavy chain CRD3 region (clone 2-100B) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996

C;Accession: PT0240

R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shans, S.; Caton, A.J.; Rovera, G.

J. Exp. Med. 173, 395-407, 1991

A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and

A;Reference number: PT0222; MUID:91108337; PMID:1899102

A;Accession: PT0240

A;Molecule type: DNA

A;Residues: 1-4 <YAM>

A;Experimental source: B lymphocyte

C;Keywords: heterotetramer; immunoglobulin

Query Match 52.4%; Score 11; DB 2; Length 4;

Best Local Similarity 66.7%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PSL 4

Db 2 PGL 4

RESULT 3

I54357
schwannin - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C:Accession: I54357
R:Huynh, D.P.; Nechiporuk, T.; Pulst, S.
Hum. Mol. Genet. 3, 1075-1079, 1994
A:Title: Alternative transcripts in the mouse neurofibromatosis type 2 (NF2) gene are co
A:Reference number: I54357; MUID:95072570; PMID:7981675
A:Accession: I54357
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-4 <RES>
A:Cross-references: GB:L28838; NID:G454836; PIDN:AAA57150.1; PID:G601923
C:Genetics: NF2
C:Gene: NF2

Query Match 42.9%; Score 9; DB 2; Length 4;
Best Local Similarity 33.3%; Pred. No. 2.8e+05;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PSL 4
|:
DB 2 PEM 4

RESULT 4
I78890
tyrosine protein kinase - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: I78890
R:Chow, L.M.; Davidson, D.; Fournel, M.; Gosselin, P.; Lemieux, S.; Lyu, M.S.; Kozak, C.
Oncogene 9, 3437-3448, 1994
A:Title: Two distinct protein isoforms are encoded by ntck, a csk-related tyrosine protei
A:Reference number: I58407; MUID:95060800; PMID:7970703
A:Accession: I78890
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3 <RES>
A:Cross-references: GB:L33339; NID:G609536; PIDN:AAA64432.1; PID:G609538
C:Genetics: p52ntk
A:Gene: p52ntk

Query Match 38.1%; Score 8; DB 3; Length 3;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PS 3
|:
DB 2 PT 3

RESULT 5
S53508
starvation-induced ribonuclease - tomato
C:Species: Lycopersicon esculentum (tomato)
C>Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 07-May-1999
C:Accession: S53508
R:Koeck, M.; Loeffler, A.; Abel, S.; Glund, K.
Plant Mol. Biol. 27, 477-485, 1995
A:Title: cDNA structure and regulatory properties of a family of starvation-induced ribo
A:Reference number: S53506; MUID:95201242; PMID:7894013
A:Accession: S53508
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-4 <ROE>

Query Match 38.1%; Score 8; DB 2; Length 4;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PS 3

I54357
schwannin - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C:Accession: I54357
R:Huynh, D.P.; Nechiporuk, T.; Pulst, S.
Hum. Mol. Genet. 3, 1075-1079, 1994
A:Title: Alternative transcripts in the mouse neurofibromatosis type 2 (NF2) gene are co
A:Reference number: I54357; MUID:95072570; PMID:7981675
A:Accession: I54357
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-4 <RES>
A:Cross-references: GB:L28838; NID:G454836; PIDN:AAA57150.1; PID:G601923
C:Genetics: NF2
C:Gene: NF2

Query Match 42.9%; Score 9; DB 2; Length 4;
Best Local Similarity 33.3%; Pred. No. 2.8e+05;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PSL 4
|:
DB 2 PEM 4

RESULT 4
I78890
tyrosine protein kinase - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: I78890
R:Chow, L.M.; Davidson, D.; Fournel, M.; Gosselin, P.; Lemieux, S.; Lyu, M.S.; Kozak, C.
Oncogene 9, 3437-3448, 1994
A:Title: Two distinct protein isoforms are encoded by ntck, a csk-related tyrosine protei
A:Reference number: I58407; MUID:95060800; PMID:7970703
A:Accession: I78890
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3 <RES>
A:Cross-references: GB:L33339; NID:G609536; PIDN:AAA64432.1; PID:G609538
C:Genetics: p52ntk
A:Gene: p52ntk

Query Match 38.1%; Score 8; DB 3; Length 3;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PS 3
|:
DB 2 PT 3

RESULT 5
S53508
starvation-induced ribonuclease - tomato
C:Species: Lycopersicon esculentum (tomato)
C>Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 07-May-1999
C:Accession: S53508
R:Koeck, M.; Loeffler, A.; Abel, S.; Glund, K.
Plant Mol. Biol. 27, 477-485, 1995
A:Title: cDNA structure and regulatory properties of a family of starvation-induced ribo
A:Reference number: S53506; MUID:95201242; PMID:7894013
A:Accession: S53508
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-4 <ROE>

Query Match 38.1%; Score 8; DB 2; Length 4;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PS 3

Db 2 PT 3
|:
RESULT 6
PT0645
T-cell receptor beta chain V-D-J region (111-1A1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0645; PT0626
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0645
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-4 <PEE>
A:Experimental source: newborn thymus, strain BALB/c, 111-1A1
A:Accession: PT0626
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-4 <PE2>
A:Experimental source: newborn thymus, strain BALB/c, 120-2CJ
C:Keywords: T-cell receptor

Query Match 38.1%; Score 8; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SL 4
|:
DB 3 SL 4

RESULT 7
RHTDRO
thyroliberin - Bombina orientalis
C:Species: Bombina orientalis
C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A90919; A01415
R:Yasuhara, T.; Nakajima, T.
Chem. Pharm. Bull. 23, 3301-3303, 1975
A:Title: Occurrence of Pyr-His-Pro-NH-2 in the frog skin.
A:Reference number: A90919; MUID:76138399; PMID:815011
A:Accession: A90919
A:Molecule type: protein
A:Residues: 1-3 <YAS>
C:Superfamily: thyroliberin precursor
C:Keywords: amidated carboxyl end; cutaneous gland; hormone; pyroglutamic acid
F.1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F.3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 33.3%; Score 7; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 P 2
|:
DB 3 P 3

RESULT 8
RHPGT
thyroliberin - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A01415
R:Nair, R.M.G.; Barrett, J.F.; Bowers, C.Y.; Schally, A.V.
Biochemistry 9, 1103-1106, 1970
A:Title: Structure of porcine thyrotropin releasing hormone.
A:Reference number: A90560; MUID:70136150; PMID:4984938
A:Accession: A01415

A:Molecule type: protein
A:Residues: 1-3 <NAI>
R:Bolex, J.; Enzmann, F.; Folkers, K.; Bowers, C.Y.; Schally, A.V.
Biochem. Biophys. Res. Commun. 37, 705-710, 1969
A:Title: The identity of chemical and hormonal properties of the thyrotropin releasing hormone
A:Reference number: A90167; MUID:70039904; PMID:4982117
A:Contents: annotation
A>Note: biological activities and Rf values (in 17 chromatographic systems) of the synthetic thyrotropin releasing hormone-like peptide - rabbit
C:Superfamily: thyroliberin precursor
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 33.3%; Score 7; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 P 2
DB 3 P 3

RESULT 9
PSSM
thyroliberin - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A93750; A01415
R:Desiderio Jr., D.M.; Burgess, R.; Dunn, T.F.; Vale, W.; Guillemin, R.; Ward, D.N.
Osg. Mass Spectrom. 5, 221-228, 1971
A:Title: The elucidation of the primary structure of the hypothalamic thyroid stimulating hormone
A:Reference number: A93750
A:Accession: A93750
A:Molecule type: protein
A:Residues: 1-3 <DES>
R:Burgus, R.; Dunn, T.F.; Desiderio, D.; Ward, D.N.; Vale, W.; Guillemin, R.
Nature 226, 321-325, 1970
A:Title: Characterization of ovine hypothalamic hypophysiotropic TSH-releasing factor.
A:Reference number: A93161; MUID:70163386; PMID:4985794
A:Contents: annotation
A>Note: physicochemical characteristics and biological activities of the natural and synthetic thyrotropin releasing hormone
C:Superfamily: thyroliberin precursor
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 33.3%; Score 7; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 P 2
DB 3 P 3

RESULT 10
A92971
thyroliberin - eastern newt (tentative sequence)
C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A92971; A01415
R:Grimm-Jorgensen, Y.; McKelvy, J.F.
J. Neurochem. 23, 471-478, 1974
A:Title: Biosynthesis of thyrotropin releasing factor by newt (Triturus viridescens) brain
A:Reference number: A92971; MUID:75035605; PMID:4214528
A:Accession: A92971
A:Molecule type: protein
A:Residues: 1-3 <GRI>
A>Note: a peptide with the chromatographic and electrophoretic characteristics of thyroliberin, or glutamic acid
C:Superfamily: thyroliberin precursor
C:Keywords: amidated carboxyl end; cutaneous gland; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 33.3%; Score 7; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 P 2
DB 3 P 3

RESULT 11
A33802
thyrotropin-releasing hormone-like peptide - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A33802
R:Cockle, S.M.; Aitken, A.; Beg, F.; Smyth, D.G.
J. Biol. Chem. 264, 7788-7791, 1989
A:Title: A novel peptide, pyroglutamylglutamylproline amide, in the rabbit prostate co-releasing factor
A:Reference number: A33802; MUID:89255196; PMID:2498305
A:Accession: A33802
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-3 <COC>
C:Superfamily: unassigned animal peptides
C:Keywords: amidated carboxyl end; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 33.3%; Score 7; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 P 2
DB 3 P 3

RESULT 12
A43391
TRH-like tripeptide - alfalfa
C:Species: Medicago sativa (alfalfa)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: A43391
R:Lackey, D.E.
J. Biol. Chem. 267, 17508-17511, 1992
A:Title: Isolation and structural determination of a novel TRH-like tripeptide, pyroglutamylglutamylproline amide, in the rabbit prostate co-releasing factor
A:Reference number: A43391; MUID:92388092; PMID:1517203
A:Accession: A43391
A:Molecule type: protein
A:Residues: 1-3 <LAC>
C:Keywords: amidated carboxyl end; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 33.3%; Score 7; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 P 2
DB 3 P 3

RESULT 13
A02147
phagocytosis-stimulating peptide (tuftsin) - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 03-Feb-1994
C:Accession: A02147
R:Nishio, K.; Constantopoulos, A.; Satoh, P.S.; Najjar, V.A.
Biochem. Biophys. Res. Commun. 47, 172-179, 1972

Search completed: March 10, 2004, 15:08:24
Job time : 21 secs

A;Title: The characteristics, isolation and synthesis of the phagocytosis stimulating pe
A;Reference number: A02147; PMID:72187087; PMID:4112769

A;Accession: A02147
A;Molecule type: protein
A;Residues: 1-4 <NIS>
A;Note: a peptide having the same structure, physical properties, and biological activit
R;Fidalgo, B.V.; Najjar, V.A.
Biochemistry 6, 3386-3392, 1967
A;Reference number: A37502; PMID:68091045; PMID:4169272
A;Contents: annotation; immunoglobulin class
C;Comment: An IgG (called leukokinin) binds reversibly to the cell membrane of neutrophil
n is essential for maximum stimulation of the phagocytic activity of neutrophils.
C;Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 33.3%; Score 7; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05; Indels 0; Gaps 0;
Matches 1; Conservative 0; Mismatches 0;

Qy 2 P 2
|
Db 3 P 3

RESULT 14

A32039
tyrosine-melanocyte-stimulating hormone release-inhibiting factor 1 - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 18-Aug-2000
C;Accession: A32039
R;Horvath, A.; Kastin, A.J.
J. Biol. Chem. 264, 2175-2179, 1989
A;Title: Isolation of tyrosine-melanocyte-stimulating hormone release-inhibiting factor
A;Reference number: A32039; PMID:89123285; PMID:2563371
A;Accession: A32039
A;Molecule type: protein
A;Residues: 1-4 <HOR>
A;Experimental source: brain
C;Superfamily: unassigned animal peptides
C;Keywords: amidated carboxyl end
F;4/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 33.3%; Score 7; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05; Indels 0; Gaps 0;
Matches 1; Conservative 0; Mismatches 0;

Qy 2 P 2
|
Db 2 P 2

RESULT 15

PL0140
carbon-monoxide dehydrogenase (EC 1.2.99.2) large chain - Pseudomonas carboxydohydrogena
C;Species: Pseudomonas carboxydohydrogena
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 28-Apr-1993
C;Accession: PL0140
R;Kraut, M.; Hugendieck, I.; Herwig, S.; Meyer, O.
Arch. Microbiol. 152, 335-341, 1989
A;Title: Homology and distribution of CO dehydrogenase structural genes in carboxydotrop
A;Reference number: PL0138; PMID:90055678; PMID:2818128
A;Accession: PL0140
A;Molecule type: protein
A;Residues: 1-4 <KRA>
C;Comment: Carbon-monoxide dehydrogenase consists of three polypeptide chains: large, me
C;Keywords: oxidoreductase

Query Match 33.3%; Score 7; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05; Indels 0; Gaps 0;
Matches 1; Conservative 0; Mismatches 0;

Qy 2 P 2
|
Db 4 P 4

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 15:01:00 ; Search time 11 Seconds
(without alignments)

18,935 Million cell updates/sec

Title: US-09-848-834A-5

Perfect score: 21

Sequence: 1 GPLS 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 18

Minimum DB seq length: 0

Maximum DB seq length: 4

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	33.3	3	1	THYL_PIG
2	7	33.3	4	1	P01151 sus scrofa
3	7	33.3	4	1	P19916 pseudomonas
4	7	33.3	4	1	P36515 saccharomyc
5	6	28.6	3	1	P01858 homo sapien
6	6	28.6	4	1	P01157 homo sapien
7	6	28.6	4	1	P35904 achatina fu
8	6	28.6	4	1	P02731 homo sapien
9	6	28.6	4	1	P58648 octopus min
10	4	19.0	4	1	P42562 hirudo medi
11	4	19.0	4	1	P42561 hirudo medi
12	4	19.0	4	1	P58707 anthopleura
13	2	9.5	3	1	P4272 vibrio fisc
14	2	9.5	4	1	P19918 pseudomonas
15	2	9.5	4	1	P42563 hirudo medi
16	2	9.5	4	1	P01162 macrocallis
17	2	9.5	4	1	P58706 anthopleura
18	1	4.8	4	1	P58705 anthopleura

ALIGNMENTS

RESULT 1
THYL_PIG
ID THYL_PIG STANDARD; PRT; 3 AA.
AC P01151;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thyroliberin (Thyrotropin releasing hormone) (TRH) (Protirelin).
OS Sus scrofa (Pig),
OS Ovis aries (Sheep),

OS Bombina orientalis (Oriental fire-bellied toad), and
OS Notopterus viridescens (Eastern newt) (Triturus viridescens).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823, 9940, 8346, 8316;
RN [1]
RP SEQUENCE.
RC SPECIES=Pig; TISSUE=Hypothalamus;
RX MEDLINE=70136150; PubMed=4984938;
RA Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;
RT "Structure of porcine thyrotropin releasing hormone.";
RL Biochemistry 9:1103-1106(1970).
RN [2]
RP SEQUENCE.
RC SPECIES=Pig;
RX MEDLINE=70039904; PubMed=4982117;
RA Boler J., Enzmann F., Folkers K., Bowers C.Y., Schally A.V.;
RT "The identity of chemical and hormonal properties of the thyrotropin
RT releasing hormone and pyroglutamyL-histidyl-proline amide.";
RL Biochem. Biophys. Res. Commun. 37:705-710(1969).
RN [3]
RP SEQUENCE.
RC SPECIES=Sheep; TISSUE=Hypothalamus;
RA Desiderio D.M. Jr., Burgus R., Dunn T.F., Vale W., Guillemin R.,
RA Ward D.N.;
RT "The elucidation of the primary structure of the hypothalamic thyroid
RT stimulating hormone releasing factor of ovine origin by means of mass
RT spectrometry.";
RL Org. Mass Spectrom. 5:221-228(1971).
RN [4]
RP SEQUENCE.
RC SPECIES=Sheep;
RX MEDLINE=70163386; PubMed=4985794;
RA Burgus R., Dunn T.F., Desiderio D.M., Ward D.N., Vale W.,
RA Guillemin R.;
RT "Characterization of ovine hypothalamic hypophysiotropic
RT TSH-releasing factor.";
RL Nature 226:321-325(1970).
RN [5]
RP SEQUENCE.
RC SPECIES=B.orientalis; TISSUE=Skin;
RX MEDLINE=76138399; PubMed=815011;
RA Yasuhara T., Nakajima T.;
RT "Letter: Occurrence of Pyr-His-Pro-NH2 in the frog skin.";
RL Chem. Pharm. Bull. 23:3301-3303(1975).
RN [6]
RP SEQUENCE.
RC SPECIES=N.viridescens;
RX MEDLINE=75035605; PubMed=4214528;
RA Grimm-Joergensen Y., McKelvy J.F.;
RT "Biosynthesis of thyrotropin releasing factor by newt (Triturus
RT viridescens) brain in vitro. Isolation and characterization of
RT thyrotropin releasing factor.";
RL J. Neurochem. 23:471-478(1974).
CC -!- FUNCTION: TRH functions as a regulator of the biosynthesis of TSH
CC in the anterior pituitary gland and as a neurotransmitter/
CC neuromodulator in the central and peripheral nervous systems.
DR PIR; A90919; RHTDQ.
DR PIR; A92971; A92971.
DR PIR; A93750; RSHST.
KW Amidation; Pyrrolidone carboxylic acid
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 3 3 AMIDATION.
SQ SEQUENCE 3 AA; 380 MW; 7761F6B000000000 CRC64;

Query Match 33.3%; Score 7; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 2 P 2

Db 3 P 3

```

RESULT 2
DCML_PSECH STANDARD; PRT; 4 AA.
ID DCML_PSECH
AC P19916;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) (CO
DE dehydrogenase subunit L) (CO-DH L) (Fragment).
DE CUTL.
GN Pseudomonas carboxydohydrogena.
OS Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae.
OX NCBI_TaxID=290;
RN [1]
RP SEQUENCE.
RX MEDLINE=90055678; PubMed=2818128;
RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT "Homology and distribution of CO dehydrogenase structural genes in
RT carboxydohydrogen bacteria.";
RL Arch. Microbiol. 152:335-341(1989).
CC -1- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
CC dioxide.
CC -1- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
CC acceptor.
CC -1- COFACTOR: Molybdenum (molybdopterin).
CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC SMALL.
CC PIR; PLO140; PLO140.
DR Oxidoreductase; Molybdenum.
KW NON_TER
FT NON_TER
SQ SEQUENCE 4 AA; 441 MW; 7761B876F00000000 CRC64;

Query Match 33.3%; Score 7; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 P 2
Db 4 P 4

RESULT 3
RM01_YEAST STANDARD; PRT; 4 AA.
ID RM01_YEAST
AC P36515;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Mitochondrial 60S ribosomal protein L1 (YmL1) (Fragment).
DE MRPL1.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE.
RX MEDLINE=91285106; PubMed=2060626;
RA Grohmann L., Graack H.-R., Kruft V., Choli T., Goldschmidt-Reisin S.,
RA Kitakawa M.;
RT "Extended N-terminal sequencing of proteins of the large ribosomal
RT subunit from yeast mitochondria.";
RL FEBS Lett. 284:51-56(1991).
DR PIR; S17255; S17255.
DR SGD; I0002681; MRPL1.
KW Ribosomal protein; Mitochondrion.
FT NON_TER
SQ SEQUENCE 4 AA; 402 MW; 7771B2D5D00000000 CRC64;

Query Match 33.3%; Score 7; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 P 2
Db 4 P 4

RESULT 4
TUFT_HUMAN STANDARD; PRT; 4 AA.
ID TUFT_HUMAN
AC P01859;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pragocytosis-stimulating peptide (Tuftsin).
DE OS Homo sapiens (Human).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=72187087; PubMed=4112769;
RA Nishioka K., Constantopoulos A., Satoh P.S., Najjar V.A.;
RT "The characteristics, isolation and synthesis of the phagocytosis
RT stimulating peptide tuftsin.";
RL Biochem. Biophys. Res. Commun. 47:172-179(1972).
RN [2]
RP IMMUNOGLOBULIN CLASS.
RX MEDLINE=68091045; PubMed=4169272;
RA Fidalgo B.V., Najjar V.A.;
RT "The physiological role of the lymphoid system. VI. The stimulatory
RT effect of leucophilic gamma globulin (leucokinin) on the phagocytic
RT activity of human polymorphonuclear leucocyte.";
RL Biochemistry 6:3386-3392(1967).
CC -1- MISCELLANEOUS: An IGG (called leucokinin) binds reversibly to the
CC cell membrane of neutrophils in the blood. Leucokininase on the
CC membrane releases the active peptide tuftsin from the gamma chain.
CC Tuftsin is essential for maximum stimulation of the phagocytic
CC activity of neutrophils.
DR PIR; A02147; A02147.
DR MIM; 191150; -.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006909; F:phagocytosis; NAS.
SQ SEQUENCE 4 AA; 501 MW; 74176321C00000000 CRC64;

Query Match 33.3%; Score 7; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 P 2
Db 3 P 3

RESULT 5
GRAM_HUMAN STANDARD; PRT; 3 AA.
ID GRAM_HUMAN
AC P01157;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE Growth-modulating peptide.
DE OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77162369; PubMed=858356;
RA Schlesinger D.H., Pickart L., Thaler M.M.;
RT "Growth-modulating serum tripeptide is glycyl-histidyl-lysine.";
RL Experientia 33:324-325(1977).
CC -1- MISCELLANEOUS: This serum tripeptide has been found to stimulate
CC growth of some cell types and to inhibit other types in vitro.

```

DR GO: GO:0001558; P:regulation of cell growth; NAS.
SQ SEQUENCE 3 AA; 340 MW; 6331E91000000000 CRC64;

Query Match 28.6%; Score 6; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 1 G 1

RESULT 6
ACH1 ACHFU STANDARD; PRT; 4 AA.
AC P35904;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Achatin-I.
OS Achatina fulica (Giant African snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Sigurethra; Achatinoidea; Achatinidae; Achatina.
OX NCBI_TaxID=6530;
RN [1]
RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
RC STRAIN=Perussac; TISSUE=Ganglion;
RX MEDLINE=9273551; PubMed=2597281;
RA Kamatani Y., Minakata H., Kenny P.T.M., Iwashita T., Watanabe K.,
RA Funase K., Sun X.P., Yongshiri A., Kim K.H., Novales-Li P.,
RA Novales E.T., Kanapi C.G., Takeuchi H., Nomoto K.,
RT "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina
RT fulica Ferussac containing a D-amino acid residue."
RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).
RN [2]
RP CHARACTERIZATION.
RC STRAIN=Perussac; TISSUE=Heart atrium;
RX MEDLINE=91264856; PubMed=1675568;
RA Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,
RA Yoshida M., Harada A., Muneoka Y., Kobayashi M.,
RT "Purification of achatin-I from the atria of the African giant snail,
RT Achatina fulica, and its possible function."
RL Biochem. Biophys. Res. Commun. 177:847-853(1991).
RN [3]
RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE=93014529; PubMed=1399265;
RA Iwashita T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,
RA "Crystal structure and molecular conformation of achatin-I
RT (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a
RT D-amino acid residue."
RL Int. J. Pept. Protein Res. 39:258-264(1992).
CC -!- FUNCTION: Neuroexcitatory peptide; increases the impulse frequency
CC and produces a spike broadening of the identified heart excitatory
CC neuron (PON); also enhances the amplitude and frequency of the
CC heart beat. Has also an effect on several other muscles.
DR PIR; A32480; A32480.
KW Hormone; D-amino acid.
FT MOD RES 2 2 D-PHENYLALANINE.
SQ SEQUENCE 4 AA; 408 MW; 6AADD9C810000000 CRC64;

Query Match 28.6%; Score 6; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 1 G 1

RESULT 7
EOSI_HUMAN
ID EOSI_HUMAN STANDARD; PRT; 4 AA.

AC P02731;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE Eosinophilotoxic peptides.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE
RX MEDLINE=76078412; PubMed=1060093;
RA Goetzi E.J., Austen K.F.;
RT "Purification and synthesis of eosinophilotoxic tetrapeptides of
RT human lung tissue: identification as eosinophil chemotactic factor of
RT anaphylaxis."
RL Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).
CC -!- MISCELLANEOUS: These peptides are released from mast cells in lung
CC (and other tissues) during hypersensitivity reactions
CC (anaphylaxis). Their activities, preferentially affecting
CC eosinophils, include chemotaxis, chemotactic deactivation, release
CC of enzymes, and stimulation of the hexose monophosphate shunt.
CC GO: GO:0006935; P:chemotaxis; IDA.
DR GO: GO:0006955; P:immune response; IDA.
DR VARIANT 1 1 V -> A (IN OTHER PEPTIDE).
FT
FT
SQ SEQUENCE 4 AA; 390 MW; 6B05B862A0000000 CRC64;
/FTID=VAR 005201.

Query Match 28.6%; Score 6; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 2 G 2

RESULT 8
OCP1_OCTMI STANDARD; PRT; 4 AA.
ID OCP1_OCTMI
AC P36648; 2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cardioactive peptides Ocp-1/Ocp-2.
OS Octopus minor (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
OX NCBI_TaxID=99766;
RN [1]
RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
RC TISSUE=Brain;
RX MEDLINE=20336815; PubMed=10876044;
RA Iwakoshi E., Hisada M., Minakata H.;
RT "Cardioactive peptides isolated from the brain of a Japanese octopus,
RT Octopus minor."
RL Peptides 21:623-630(2000).
CC -!- FUNCTION: Cardioactive; has both positive chronotropic and
CC inotropic effects on the heart. Ocp-2 is a 1000 time less
CC active than Ocp-1.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: Ocp-2 has L-Phe instead of D-Phe.
CC -!- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.
KW Hormone; D-amino acid.
FT MOD RES 2 2 D-PHENYLALANINE (IN OCP-1).
SQ SEQUENCE 4 AA; 394 MW; 6AA879C810000000 CRC64;

Query Match 28.6%; Score 6; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G 1
DB 1 G 1

Db 1 G 1

RESULT 9

OCPI OCTMI STANDARD; PRT; 4 AA.
 AC P58643;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cardioactive peptides Ocp-3/Ocp-4.
 OS Octopus minor (Octopus).
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
 OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
 OX NCBI_TaxID=89766;
 RN [1]
 RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
 RC TISSUE=Brain;
 RX MEDLINE=20336815; PubMed=10876044;
 RA Iwakoshi E., Hsada M., Minakata H.;
 RT "Cardioactive peptides isolated from the brain of a Japanese octopus,
 Octopus minor";
 RL Peptides 21:623-630(2000).
 CC -!- FUNCTION: Cardioactive; has both positive chronotropic and
 inotropic effects on the heart. Ocp-4 is a 1000 time less
 active than Ocp-3.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PFM: Ocp-4 has D-Ser instead of L-Ser.
 CC -!- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.
 KW Hormone; D-amino acid.
 FT MOD RES 2 D-SERINE (IN OCP-4).
 SQ SEQUENCE 4 AA; 463 MW; 6AB365B810000000 CRC64;

Query Match 28.6%; Score 5; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.4e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 G 1

Db 1 G 1

RESULT 10

FAR3_HIRME STANDARD; PRT; 4 AA.
 AC P42562;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE FMRamide-like neuropeptide YURF-amide.
 OS Hirudo medicinalis (Medicinal leech).
 OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
 OC Arhynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
 OX NCBI_TaxID=6421;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92195954; PubMed=1686933;
 RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
 RT "Identification of Rfamide neuropeptides in the medicinal leech."
 RL Peptides 12:897-908(1991).
 CC -!- SIMILARITY: Belongs to the RARP (FMRamide related peptide)
 family.
 CC Neuropeptide; Amidation.
 KW Neuropeptide; Amidation.
 FT MOD RES 4
 SQ SEQUENCE 4 AA; 598 MW; 69D4073B30000000 CRC64;

Query Match 19.0%; Score 4; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.4e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 L 4

Db 2 L 2

RESULT 11

FLRF_HIRME STANDARD; PRT; 4 AA.
 AC P42561;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE FLRFamide.
 OS Hirudo medicinalis (Medicinal leech), and
 OS Helisoma trivolvis (Snail).
 OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
 OC Arhynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
 OX NCBI_TaxID=6421, 27815;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=H. medicinalis;
 RX MEDLINE=92195954; PubMed=1686933;
 RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
 RT "Identification of Rfamide neuropeptides in the medicinal leech."
 RL Peptides 12:897-908(1991).
 RN [2]

SEQUENCE.
 RP SPECIES=H. trivolvis; TISSUE=Kidney;
 RX MEDLINE=94286417; PubMed=7912428;
 RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
 RT "FMRamide-related peptides from the kidney of the snail, Helisoma
 trivolvis";
 RL Peptides 15:31-36(1994).
 CC -!- SIMILARITY: Belongs to the RARP (FMRamide related peptide)
 family.
 CC Neuropeptide; Amidation.
 FT MOD RES 4
 SQ SEQUENCE 4 AA; 582 MW; 69D40729A0000000 CRC64;

Query Match 19.0%; Score 4; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.4e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 L 4

Db 2 L 2

RESULT 12

FLRN ANTEL STANDARD; PRT; 4 AA.
 AC P58707;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Antho-RNamide.
 OS Anthopleura elegantissima (Sea anemone).
 OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
 OC Nynanthae; Actiniidae; Anthopleura.
 OX NCBI_TaxID=6110;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RX MEDLINE=90319122; PubMed=1373541;
 RA Grimelikhuijzen C.J.P., Rinehart K.L. Jr., Jacob E., Graff D.,
 RA Reinscheid R.K., Nothacker H.-P., Staley A.L.;
 RT "Isolation of L-3-phenyllactyl-Leu-arg-Asn-NH2 (Antho-RNamide), a sea
 anemone neuropeptide containing an unusual amino-terminal blocking
 group";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5410-5414(1990).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Neuron specific.
 CC -!- MASS SPECTROMETRY: MW=549.3; METHOD=FAB.
 DR PIR: A35779; A35779.
 KW Neuropeptide; Amidation.
 FT MOD RES 4
 SQ SEQUENCE 4 AA; 598 MW; 69D4073B30000000 CRC64;

Query Match 19.0%; Score 4; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.4e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 L 4

Db 2 L 2

```
SQ SEQUENCE 4 AA; 549 MW; 64540729A000000000 CRC64;
Query Match 19.0%; Score 4; DB 1; Length 4;
Best Local Similarity 100.0%; Pred.No. 1.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 L 4
Db 2 L 2

RESULT 13
LUXE_VIBFI STANDARD; PRT; 3 AA.
AC P24272;
DT 01-NAR-1992 (Rel. 21, Created)
DT 01-NAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-1995 (Rel. 34, Last annotation update)
DE Long-chain-fatty-acid--luciferin-component ligase (EC 6.2.1.19) (Acyl-
protein synthetase) (Fragment).
GN LUXE.
OS Vibrio fischeri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91072226; PubMed=2254256;
RA Swartzman E., Kapoor S., Graham A.F., Meighan E.A.;
RT "A new Vibrio fischeri lux gene precedes a bidirectional termination
site for the lux operon."
RL J. Bacteriol. 172:6797-6802(1990).
CC -!- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.
IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE
FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS
SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.
CC -!- CATALYTIC ACTIVITY: ATP + an acid + protein = AMP + diphosphate +
an acyl-protein thioester.
CC -!- PATHWAY: Bioluminescent fatty acid reduction system; second step.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; M62812; -; NOT_ANNOTATED_CDS.
KW Luminescence; Ligase.
FT NON TER 1 1
SQ SEQUENCE 3 AA; 374 MW; 6AA330300000000000 CRC64;

Query Match 9.5%; Score 2; DB 1; Length 3;
Best Local Similarity 0.0%; Pred.No. 1.4e+05;
Matches 0; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 L 4
Db 1 I 1

RESULT 14
DCMS_PSECH STANDARD; PRT; 4 AA.
AC P19918;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carbon monoxide dehydrogenase small chain (EC 1.2.99.2) (CO
dehydrogenase subunit S) (CO-DH S) (Fragment).
GN CUTS.
OS Pseudomonas carboxydohydrogena.
```

```
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae.
OX NCBI_TaxID=290;
RN [1]
RP SEQUENCE.
RX MEDLINE=90055678; PubMed=2818128;
RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT "Homology and distribution of CO dehydrogenase structural genes in
carboxydohydrogenic bacteria."
RL Arch. Microbiol. 152:335-341(1989).
CC -!- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
dioxide.
CC -!- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
acceptor.
CC -!- COFACTOR: Binds 2 2Fe-2S clusters.
CC -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
SMALL.
DR PIR; P10146; P10146.
KW Oxidoreductase; Metal-binding; Iron-sulfur; Iron; 2Fe-2S.
FT NON TER 4 4
SQ SEQUENCE 4 AA; 420 MW; 6DD33DD6F000000000 CRC64;

Query Match 9.5%; Score 2; DB 1; Length 4;
Best Local Similarity 0.0%; Pred.No. 1.4e+05;
Matches 0; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 L 4
Db 1 M 1

RESULT 15
PAR4_HIRME STANDARD; PRT; 4 AA.
AC P42583;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRamide-like neuropeptide YMRP-amide.
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Achiynobdellida; Hirudiniiformes; Hirudinidae; Hirudo.
OX NCBI_TaxID=6421;
RN [1]
RP SEQUENCE.
RX MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech."
RL Peptides 12:897-908(1991).
CC -!- SIMILARITY: Belongs to the FARP (FMRamide related peptide)
family.
KW Neuropeptide; Amidation.
FT MOD RES 4 4
SQ SEQUENCE 4 AA; 616 MW; 69D4068B3000000000 CRC64;

Query Match 9.5%; Score 2; DB 1; Length 4;
Best Local Similarity 0.0%; Pred.No. 1.4e+05;
Matches 0; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 L 4
Db 2 M 2

Search completed: March 10, 2004, 15:06:59
Job time : 11 secs
```

	01-JUN-2003 (TREMBLrel. 24, Last sequence update)	01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)	01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE	Neuropeptide Gwa.	

Db 2 L 2

```
RESULT 3
Q08433 PRELIMINARY; PRT; 4 AA.
AC Q08433;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Bilirubin UDP-glucuronosyltransferase (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Gunn;
RX MEDLINE=91282759; PubMed=1840486;
RA Sato H., Aono S., Kashiwamata S., Koiwai O.;
RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the
XT hyperbilirubinemic Gunn rat."; 177:1161-1164 (1991).
RL Biochem. Biophys. Res. Commun. 177:1161-1164 (1991).
DR EMBL; S38636; AAB19259.1; -.
DR GO; GO:0016740; F:transferase activity; IEA.
KW Transferase.
FT NON TER 1
SQ SEQUENCE 4 AA; 473 MW; 633732C420000000 CRC64;

Query Match 19.0%; Score 4; DB 11; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 L 4
Db 3 L 3
```

Search completed: March 10, 2004, 15:07:51
Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 15:00:25 ; Search time 52 Seconds
(without alignments)
21.734 Million cell updates/sec

Title: US-09-848-834A-5

Perfect score: 21

Sequence: 1 GPLS 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 15518

Minimum DB seq length: 0

Maximum DB seq length: 4

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Genesep29Jan04:*

1: Genesep1980s:*

2: Genesep1990s:*

3: Genesep2000s:*

4: Genesep2001s:*

5: Genesep2002s:*

6: Genesep2003as:*

7: Genesep2003bs:*

8: Genesep2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	4	4 AAB68641	Aab68641 Peptide 1
2	21	100.0	4	5 AAU11416	AAU11416 Synthetic
3	18	85.7	4	2 AAU03253	AAU03253 Peptidyl-
4	16	76.2	4	2 AAR32384	Aar32384 Fibrinoge
5	16	76.2	4	2 AAR49798	Aar49798 Sequence
6	16	76.2	4	5 ABG31105	ABG31105 Linker pe
7	15	71.4	4	2 AAU03252	AAU03252 Peptidyl-
8	15	71.4	4	2 AAU52443	AAU52443 Loop regi
9	15	71.4	4	2 AAU03640	AAU03640 Antimicro
10	15	71.4	4	3 AAY80113	Aay80113 Proteasom
11	15	71.4	4	3 AAB30516	Aab30516 A peptide
12	15	71.4	4	5 ABG70996	ABG70996 Synthetic
13	15	71.4	4	5 AAU74938	AAU74938 NF-kappaB
14	15	71.4	4	5 ABG31209	ABG31209 Rat delta
15	15	71.4	4	5 AAU29577	AAU29577 Metallope
16	15	71.4	4	5 AAU29607	AAU29607 Metallope
17	15	71.4	4	5 ABG31992	ABG31992 Proteasom
18	15	71.4	4	6 ABG75753	ABG75753 Proteasom
19	14	66.7	4	2 AAR31245	Aar31245 HIV princ
20	14	66.7	4	2 AAR30051	Aar30051 Conserved
21	14	66.7	4	2 AAR26680	Aar26680 HIV-PND-p
22	14	66.7	4	2 AAR46184	Aar46184 Serine pr
23	14	66.7	4	2 AAR45996	Aar45996 Serine pr
24	14	66.7	4	2 AAR45975	Aar45975 Serine pr
25	14	66.7	4	2 AAR46017	Aar46017 Serine pr

26	14	66.7	4	2 AAR45954	Aar45954 Serine pr
27	14	66.7	4	2 AAR46038	Aar46038 Serine pr
28	14	66.7	4	2 AAR46205	Aar46205 Serine pr
29	14	66.7	4	2 AAU04997	AAU04997 Immunomod
30	14	66.7	4	2 AAU04964	AAU04964 Immunomod
31	14	66.7	4	2 AAU41982	AAU41982 Immunomod
32	14	66.7	4	2 AAU45045	AAU45045 Immunomod
33	14	66.7	4	2 AAU17689	AAU17689 Substrate
34	14	66.7	4	2 AAU03256	AAU03256 Peptidyl-
35	14	66.7	4	2 AAU52391	AAU52391 Beta-turn
36	14	66.7	4	2 AAU51152	AAU51152 Cysteine
37	14	66.7	4	2 AAU51141	AAU51141 Methionin
38	14	66.7	4	2 AAU90952	AAU90952 Collagen
39	14	66.7	4	2 AAU51990	AAU51990 Peptide h
40	14	66.7	4	2 AAU51180	AAU51180 Peptide c
41	14	66.7	4	2 AAU39839	AAU39839 Elastin p
42	14	66.7	4	2 AAU09588	AAU09588 Antimicro
43	14	66.7	4	2 AAU09408	AAU09408 Immunoact
44	14	66.7	4	2 AAU09457	AAU09457 Immunoact
45	14	66.7	4	3 AAU96519	AAU96519 Linker 4

ALIGNMENTS

RESULT 1

AAB68641

ID AAB68641 standard; peptide; 4 AA.

XX

AC AAB68641;

XX

DT 30-APR-2001 (first entry)

XX

DE Peptide linker.

XX

KW Cytostatic; immune response; HER-2; human; epitope; cancer; breast;
ovarian; lung; prostate; colon.

XX

OS Unidentified.

XX

PN WC200108636-A2.

XX

PD 08-FEB-2001.

XX

PF 03-AUG-2000; 2000WQ-US021222.

XX

PR 03-AUG-1999; 99US-0146869P.

XX

XX (OHIS) UNIV OHIO STATE.

PA

PI Kaumaya PT, Stevens VC, Triozzi PL;

XX

XX WPI; 2001-182849/18.

XX

PT Compositions comprising polypeptides and polynucleotides for stimulating
the immune system and for treating malignancies associated with
overexpression of the HER-2 protein.

XX

PS Claim 5; Page 38; 51pp; English.

XX

CC The present invention relates to compositions for stimulating the immune
system and for treating malignancies associated with overexpression of
the HER-2 protein. The compositions comprise immunogenic groups of the
HER-2 proteins. The present sequence is one such peptide used in the
compositions of the present invention. The compositions can be used for
treating cancer, e.g. breast, ovarian, lung, prostate and colon cancers

XX

SQ Sequence 4 AA;

Query Match

Best Local Similarity 100.0%; Score 21; DB 4; Length 4;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 GPSL 4
Db      1 GPSL 4

RESULT 2
AAU11416
ID AAU11416 standard; peptide; 4 AA.
XX
AC AAU11416;
XX
DT 12-MAR-2002 (first entry)
XX
DE Synthetic spacer peptide #1.
XX
KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer;
KW spacer peptide.
XX
OS Synthetic.
XX
DN W0200185763-A2.
XX
PD 15-NOV-2001.
XX
PF 04-MAY-2001; 2001WO-US014363.
XX
PR 05-MAY-2000; 2000US-0202328P.
XX
PA (APHT-) APHTON CORP.
XX
PI Grimes S, Michaeli D, Stevens VC;
XX WPI; 2002-049440/06.
XX
DR Novel synthetic immunogen for inducing immune response against
PT Gonadotropin releasing hormone, comprises fusion peptide having
PT promiscuous helper T-cell peptide epitope and immunomic peptide epitope
PT or its analog.
XX
PS Claim 10; Page 6; 43pp; English.
XX
CC The invention relates to a synthetic immunogen for inducing specific
CC antibodies against gonadotropin releasing hormone (GnRH) also known as
CC luteinising hormone releasing hormone (LHRH) comprising a fusion peptide
CC which comprises a promiscuous helper T-cell peptide epitope and
CC immunomic peptide epitope or its analogue. The synthetic immunogen is
CC useful inducing an immune response against GnRH in an animal subject, and
CC as such is useful as a contraceptive and in the treatment of diseases
CC such as cancer (of the breast, uterus and other gynaecological cancer),
CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
CC prostate cancer. The immunogen is effective in eliciting high and
CC specific anti-GnRH antibody titres. The present sequence is a synthetic
CC spacer peptide used in the immunogen of the invention
XX
SQ Sequence 4 AA;
Query Match      100.0%; Score 21; DB 5; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GPSL 4
Db      1 GPSL 4

RESULT 3
AAW03253
ID AAW03253 standard; peptide; 4 AA.
XX

AC AAW03253;
XX
DT 16-MAR-1997 (first entry)
XX
DE Peptidyl-aldehyde inhibitor of MPC proteolytic activity.
XX
KW BrAAP; inhibitor; branched chain amino acid protease; MPC;
KW multicatalytic proteinase complex; intracellular proteolysis; mitosis;
KW cachexia; autoimmune diseases; transplant rejection; peptidyl aldehyde;
KW proteasome.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "benzyloxycarbonyl-Gly"
FT Modified-site 4 /note= "Leu-CHO; i.e. the terminal COOH group is replaced
FT by an aldehyde group"
XX
PN US5580854-A.
XX
PD 03-DEC-1996.
XX
PF 04-MAR-1994; 94US-00206789.
XX
PR 02-DEC-1991; 91US-00801324.
PR 26-MAY-1993; 93US-00068273.
PR 02-JUN-1994; 94US-00253185.
XX
PA (MOUN ) MOUNT SINAI SCHOOL MEDICINE.
XX
PI Vinitzky A, Orlowski M, Cardozo C;
XX WPI; 1997-033590/03.
XX
DR New peptidyl aldehyde cpds. - are inhibitors of the BrAAP component of
PT the multi-catalytic proteinase complex.
XX
PS Claim 7; Col 6; 16pp; English.
XX
CC New peptidyl aldehydes are provided which are of general formula 2-P4-P3-
CC P2-P1-CHO, in which P1-CHO is reduced Leu, Val or Ile, P2 is Phe or Ala,
CC P3 is Pro or hydroxy-Pro, P4 is Gly, and Z is benzyloxy- carbonyl,
CC benzoylglycine, tert-butoxycarbonyl or acetyl. The compounds are
CC inhibitors of the BrAAP component of the multicatalytic proteinase
CC complex (MPC). They are thus useful for inhibiting undesired
CC intracellular proteolysis and mitosis in excessively proliferating cells,
CC such as in cachexia, and inhibiting unwanted immune response, such as in
CC autoimmune diseases and transplant rejection. The present sequence
CC represents one of three preferred peptides fitting the above general
XX formula
SQ Sequence 4 AA;
Query Match      85.7%; Score 18; DB 2; Length 4;
Best Local Similarity 75.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GPSL 4
Db      1 GPSL 4

RESULT 4
AAR32384
ID AAR32384 standard; peptide; 4 AA.
XX
AC AAR32384;
XX
DT 25-MAR-2003 (revised)
DT 01-JUL-1993 (first entry)
XX

```

DE Fibrinogen binding peptide generated by scrambling.

XX platelet aggregation; fibrinogen A-alpha chain; protein scrambling;
KW GPR-like peptide.

XX Synthetic.

XX WO9304079-A1.

XX PD 04-MAR-1993.

XX PF 20-AUG-1992; 92WO-US006933.

XX PR 21-AUG-1991; 91US-00813315.

XX PA (RECE-) RECEPTOR LAB INC.

XX PI Venton DL, Hopfinger AJ, Le Breton G;

XX DR WPI; 1993-093932/11.

XX PT Identifying peptide(s) which bond to predetermined targets - by random
PT degradation and recombination of peptide(s) and isolating bound
PT peptide(s).

XX PS Example 5; Page 60; 89pp; English.

XX CC Synthetic peptides beginning with the sequence Gly-Pro-Arg will bind to
CC fibrinogen and prevent fibrin polymerisation. A scrambling reaction was
CC used to determine whether unique GPR-like peptides could be created from
CC a reaction seeded with GPR and if these new products could bind to
CC fibrinogen. The peptide GPRL was one of the new peptides to be generated
CC by hydrolysis/synthesis of fibrinogen/GPR. It inhibits fibrin
CC polymerisation but its affinity was less than that of GPR or GPRP.
CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 4 AA;

Query Match 76.2%; Score 16; DB 2; Length 4;
Best Local Similarity 75.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GPSL 4
DB |||
1 GPRL 4

RESULT 5

AAR49798

XX AAR49798 standard; peptide; 4 AA.

XX AC AAR49798;

XX DT 25-MAR-2003 (revised)

XX DT 23-AUG-1994 (first entry)

XX DE Sequence of peptide which binds to human fibrinogen.

XX KW Random degradation; recombination; scrambling reaction.

XX OS Synthetic.

XX PN WO9404558-A1.

XX PD 03-MAR-1994.

XX PF 09-AUG-1993; 93WO-US008231.

XX PR 21-AUG-1992; 92US-00932200.

XX PA (RECE-) RECEPTOR LAB INC.

XX PI Venton DL, Hopfinger AJ;

XX WPI; 1994-083103/10.

XX PT Identifying peptides which binds to a specific target - by contacting
PT target with scrambled equilibrium mixt. of many peptide derived from
PT protein by incubation with protease, for detecting potential therapeutic
PT agents.

XX PS Example; Page 54; 97pp; English.

XX CC The inventors claim a method for inexpensively and rapidly producing a
CC large and varied population of peptides and screening this varied
CC population for the presence of peptides which bind to a target, for
CC example, a macromolecule associated with a particular physiological
CC function. The specific binding peptides are isolated and sequenced,
CC synthesised on a large-scale, their biological activity is demonstrated,
CC and then subjected to clinical testing. The random population of peptides
CC is generated by employing a scrambling system which utilises one or more
CC proteases, esp. papain, pepsin, bromelain, thermolysin, trypsin, pronase,
CC chymotrypsin, subtilisin and dipeptidyl peptidase IV. A typical starting
CC protein is casein. Targets are esp. receptors involved in physiological
CC processes, partic. fibrinogen; sickle cell haemoglobin; collagenase IV;
CC rennin; Gp. IIb/IIIa or phospholipase A2. It has been demonstrated that
CC synthetic peptides beginning with the sequence GPR will bind to
CC fibrinogen. Such peptides include GPR, GPRP and GPRV. HPLC fractions
CC bound to fibrinogen were screened on CE. Three of these GPR peaks were
CC collected and sequenced and found to be GPRL, GPRP and LPRK. In addition,
CC a GPR-independent peak was also collected and found to have the sequence
CC DKPDNF. (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 4 AA;

Query Match 76.2%; Score 16; DB 2; Length 4;
Best Local Similarity 75.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GPSL 4
DB |||
1 GPRL 4

RESULT 6

ABG31105

XX ID ABG31105 standard; peptide; 4 AA.

XX AC ABG31105;

XX DT 21-OCT-2002 (first entry)

XX DE Linker peptide for glucocerebrosidase fusion proteins #10.

XX KW Glucocerebrosidase; GCR; fusion protein; IGG1; linker peptide; Fc;

XX KW glycolipid storage disorder; Gaucher's disease; Fabry's disease;

XX OS Tay-Sachs disease; nephrotropic; antilipemic.

XX OS Synthetic.

XX PN WO200257435-A2.

XX PD 25-JUL-2002.

XX PF 27-DEC-2001; 2001WO-EP015328.

XX PR 18-JAN-2001; 2001EP-00101056.

XX PA (MERE) MERCK PATENT GMBH.

XX PI Schumacher S, Gillies S;

XX DR WPI; 2002-599719/64.

XX PT New fusion proteins with glucocerebrosidase activity, useful for treating
PT glycolipid storage disorders, specifically Gaucher's disease, Fabry's

PT New peptidyl aldehyde cpds. - are inhibitors of the BraAP component of
 PT the multi:catalytic proteinase complex.
 XX
 PS Claim 6; Col 6; 16pp; English.
 XX
 CC New peptidyl aldehydes are provided which are of general formula Z-P4-P3-
 CC P2-PI-CHO, in which P1-CHO is reduced Leu, Val or Ile, P2 is Phe or Ala,
 CC P3 is Pro or hydroxy-Pro, P4 is Gly, and Z is benzyl- or carbonyl,
 CC benzoylglycine, tert-butoxycarbonyl or acetyl. The compounds are
 CC inhibitors of the BraAP component of the multicatalytic proteinase
 CC complex (MPC). They are thus useful for inhibiting undesired
 CC intracellular proteolysis and mitosis in excessively proliferating cells,
 CC such as in cachexia, and inhibiting unwanted immune response, such as in
 CC autoimmune diseases and transplant rejection. The present sequence
 CC represents one of three preferred peptides fitting the above general
 CC formula
 XX
 SQ Sequence 4 AA;
 Query Match 71.4%; Score 15; DB 2; Length 4;
 Best Local Similarity 75.0%; Pred. NO. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GPGL 4
 Db 1 GPFL 4
 RESULT 8
 AAW52443
 ID AAW52443 standard; peptide; 4 AA.
 AC AAW52443;
 XX
 XX 01-JUL-1998 (first entry)
 DE Loop region used in cyclic peptide of the invention.
 XX
 XX Loop region; cyclic peptide; antimicrobial; disinfectant; therapy;
 KW Preservative; amphipathic anti-parallel beta-sheet region; plant disease.
 XX
 OS Synthetic.
 XX
 XX WO9803192-A1.
 XX
 PD 29-JAN-1998.
 XX
 XX 23-JUL-1997; 97WO-US012974.
 XX
 XX 24-JUL-1996; 96US-00685389.
 XX
 XX (INTR-) INTRABIOTICS PHARM INC.
 XX
 XX Chang C, Gu L, Chen J;
 XX
 XX WPI; 1998-120472/11.
 DR
 XX
 XX New cyclic peptide(s) with antimicrobial activity - contain amphipathic
 PT beta-sheet, loop and beta-turn regions, have better activity,
 PT bioavailability and protease resistance than linear analogues.
 XX
 XX Claim 4; Page 151; 160pp; English.
 XX
 CC This sequence represents a loop region used in a peptide of the
 CC invention. The peptides are cyclic peptides (I), which have: (a) an
 CC amphipathic anti-parallel beta-sheet region (SR), a loop region (LR) and
 CC a beta-turn region (TR); (b) a net positive charge at physiological pH;
 CC and (c) at least one basic amino acid (aa) in LR or TR. (I) are broad
 CC spectrum antimicrobials, specifically for use against E. coli,
 CC Pseudomonas aeruginosa, methicillin-resistant Staphylococcus aureus
 CC (MRSA), vancomycin-resistant Enterococcus faecium and penicillin-
 CC resistant Streptococcus pneumoniae. More generally they are active
 CC against Gram-positive or -negative bacteria, fungi, yeast and protozoa.

PT disease or Tay-Sachs disease.
 XX
 PS Disclosure; Page 9; 25pp; English.
 XX
 CC The invention relates to a fusion protein, which comprises essentially of
 CC an immunoglobulin molecule (Ig e.g. IgG1 Fc or IgG2 Fc) or its fragment
 CC and a non-immunoglobulin molecule. The non-immunoglobulin molecule is a
 CC protein having the biological activity of glucocerebrosidase (GCR-like
 CC protein). Also included are a DNA sequence encoding the fusion protein
 CC comprising: (a) a signal/leader sequence; (b) an Ig molecule; or (c) a
 CC target protein sequence having the biological activity of GCR. Also
 CC included are an expression vector comprising the fusion protein DNA, a
 CC host cell for expressing the novel fusion protein and comprising the
 CC vector, producing the novel fusion protein and a pharmaceutical
 CC composition comprising the novel fusion protein, and at least one
 CC pharmaceutical carrier, diluent or excipient. The fusion protein is
 CC useful for treating glycolipid storage disorders, specifically Gaucher's
 CC disease, Fabry's disease or Tay-Sachs disease. The fusion protein is also
 CC useful for the manufacture of a pharmaceutical composition for treating
 CC these diseases. The present sequence represents a linker peptide suitable
 CC for inclusion in a fusion protein of the invention
 XX
 SQ Sequence 4 AA;
 Query Match 76.2%; Score 16; DB 5; Length 4;
 Best Local Similarity 75.0%; Pred. NO. 1.4e+06;
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 GPGL 4
 Db 1 GPRL 4
 RESULT 7
 AAW03252
 ID AAW03252 standard; peptide; 4 AA.
 AC AAW03252;
 XX
 XX 16-MAR-1997 (first entry)
 DE Peptidyl-aldehyde inhibitor of MPC proteolytic activity.
 XX
 XX BraAP; inhibitor; branched chain amino acid protease; MPC;
 KW multicatalytic proteinase complex; intracellular proteolysis; mitosis;
 KW cachexia; autoimmune diseases; transplant rejection; peptidyl aldehyde;
 KW proteasome.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT /note= "benzyloxycarbonyl-Gly"
 FT Modified-site 4
 FT /note= "Leu-CHO; i.e. the terminal COOH group is replaced
 FT by an aldehyde group"
 XX
 XX US5580854-A.
 PN
 XX
 PD 03-DEC-1996.
 XX
 XX 04-MAR-1994; 94US-00206789.
 XX
 XX 02-DEC-1991; 91US-00801324.
 PR 26-MAY-1993; 93US-00068273.
 PR 02-JUN-1994; 94US-00253185.
 XX
 XX (MOUN) MOUNT SINAI SCHOOL MEDICINE.
 XX
 XX Vinitsky A, Orłowski M, Cardozo C;
 PI
 XX WPI; 1997-033590/03.
 DR
 XX

CC Apart from clinical uses, (I) are also used as disinfectants and
 CC preservatives for medical equipment, foods, cosmetics etc., also for
 CC treatment of plant diseases. Compared with non-cyclised analogues (i.e.
 CC tachyplesin and protegrin type peptides), (I) and are more effective,
 CC with better bioavailability and/or serum half-life (increased resistance
 CC to proteolysis). They are more suitable for oral administration, can be
 CC used at lower doses and are unlikely to induce development of resistant
 CC strains
 XX
 SQ Sequence 4 AA;

Query Match 71.4%; Score 15; DB 2; Length 4;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GPGL 4
 DB 1 GPQI 4

RESULT 9
 AAY09640
 ID AAY09640 standard; peptide; 4 AA.

XX AC AAY09640;
 XX DT 21-JUL-1999 (first entry)

XX DE Antimicrobial cyclic peptide loop region SEQ ID NO:88.

XX KW Cyclic peptide; antimicrobial; amphiphilic beta-sheet; biostatic;
 KW bicydial; bacteria; yeast; Candida species; fungi; protozoa;
 KW Escherichia coli; Pseudomonas aeruginosa; infection; preservative;
 KW vancomycin-resistant Enterococcus faecium; disinfectant; food;
 KW methicillin-resistant Staphylococcus aureus; medical equipment;
 KW penicillin-resistant Streptococcus pneumoniae; cosmetic.

XX OS Synthetic.

XX PN WO9921879-A1.

XX PD 06-MAY-1999.

XX PF 27-OCT-1997; 97WO-US019557.

XX PR 27-OCT-1997; 97WO-US019557.

XX PA (INTR-) INTRABIONICS PHARM INC.

XX PI Chang C, Gu L, Chen J;

XX DR WPI; 1999-312941/26.

XX PT Antimicrobial cyclic peptides with amphiphilic beta-sheet region.

XX PS Claim 4; Page 152; 167pp; English.

XX CC The present invention describes cyclic peptides (I): (i) comprising an
 CC amphiphilic anti-parallel beta-sheet region (A), a loop region (B) and a
 CC beta-turn (C); (ii) having net positive charge at physiological pH; and
 CC (iii) including at least one basic amino acid (aa) in (B) or (C).
 CC AAY09554 to AAY09683 and AAY17301 to AAY17385 represent specifically
 CC claimed cyclic peptides and peptide segments. (I) are broad-spectrum
 CC antimicrobials (biostatic or biocidal), effective against bacteria, yeast
 CC (e.g. Candida species), fungi and protozoa. Particularly they are used to
 CC control Escherichia coli, Pseudomonas aeruginosa, vancomycin-resistant
 CC Enterococcus faecium, methicillin-resistant Staphylococcus aureus and
 CC penicillin-resistant Streptococcus pneumoniae. (A) are used to treat or
 CC prevent infections in animals or plants, also as preservatives and
 CC disinfectants for medical equipment, foods, cosmetics, optionally as
 CC mixtures or in combination with other antimicrobials. Compared with
 CC linear analogues, (I) are more effective with better bioavailability
 CC and/or serum half-life (better resistance to proteolysis, allowing lower

CC doses and making them more suitable for oral delivery). Since (I) are
 CC structurally related to naturally occurring antimicrobial peptides, they
 CC are less likely to induce development of resistant strains
 XX
 SQ Sequence 4 AA;

Query Match 71.4%; Score 15; DB 2; Length 4;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GPGL 4
 DB 1 GPQI 4

RESULT 10
 AAY80113
 ID AAY80113 standard; peptide; 4 AA.

XX AC AAY80113;

XX DT 17-MAY-2000 (first entry)

XX DE Proteasome inhibitor peptide #1.

XX KW Proteasome inhibitor; NF-kappa B; bone morphogenic protein; BMP;
 KW bone growth; hair growth; osteoporosis; hyperparathyroidism;
 KW periodontal disease; metastatic bone disease; osteolytic bone disease;
 KW osteopathic; vulnary; antimetastatic; nuclear factor kappa B.

XX OS Unidentified.

XX FH Key Location/Qualifiers

XX FT Modified-site 1

XX FT Modified-site 4 /note= "Gly is modified to N-carboxybenzoyl-Gly"

XX FT Modified-site 4 /note= "Leu is modified to Leu-CHO"

XX PN WO200002548-A2.

XX PD 20-JAN-2000.

XX PF 09-JUL-1999; 99WO-US015533.

XX PR 10-JUL-1998; 98US-00113947.

XX PA (OSTE-) OSTEOSCREEN.

XX PI Mundy GR, Garrett IR, Rossini G;

XX DR WPI; 2000-171065/15.

XX PT Compound that inhibits the activity of NF-kappa B useful for enhancing
 XX bone formation.

XX PS Disclosure; Page 24; 39pp; English.

XX CC A method has been developed for enhancing bone formation, treating
 CC pathological dental conditions, and treating degenerative joint
 CC conditions by administration a compound that inhibits the activity of NF-
 CC kappa B or that inhibits proteasomal activity or that inhibits production of
 CC proteasome proteins. The method can be used for enhancing bone formation,
 CC treating pathological dental conditions, degenerative bone conditions,
 CC osteoporosis, bone fracture or deficiency, primary or secondary
 CC hyperparathyroidism, periodontal disease or defect, metastatic bone
 CC disease, osteolytic bone disease, post-plastic surgery, post-prosthetic
 CC joint surgery, and post-dental implantation, and for stimulating hair
 CC growth. The compounds may also be useful in wound healing or tissue
 CC repair. The present sequence represents a proteasome inhibitor peptide
 XX Sequence 4 AA;

Query Match

71.4%; Score 15; DB 3; Length 4;

Best Local Similarity 75.0%; Pred. No. 1.4e+06; Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GPSL 4
Db 1 GPFL 4

RESULT 11
AAB30516
ID AAB30516 standard; peptide; 4 AA.
AC AAB30516;
XX
XX
DT 06-MAR-2001 (first entry)
DE
XX
XX
A peptide which inhibits proteasomal activity or NF-kappaB.
KW Transcription factor NF-kB; myeloma bone disease; proteasomal enzyme;
KW multiple myeloma; osteopenia; osteolytic lesion; osteopetrosis; Ibg2b;
KW bone fracture; osteolytic bone disease; myeloma bone disease.
XX
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-carbobenzoyl attached"
FT
FT
XX WO2000061167-A2.
XX
XX 19-OCT-2000.
XX
XX 07-APR-2000; 2000WO-US009121.
XX
XX 09-APR-1999; 99US-00289229.
XX
XX (OSTE-) OSTEOSCREEN INC.
XX
XX Mundy GR;
XX
XX WPI; 2000-686989/67.
XX
XX Identifying a compound effective in treating multiple myeloma and myeloma bone disease, involves subjecting the compound to an assay determining its ability to inhibit NF-kB or proteasomal activity.
XX
XX Disclosure; Page 7; 22pp; English.
XX
XX AAB30515-17 represent exemplary compounds of the invention which inhibit proteasomal activity or the transcription factor NF-kB. The specification describes a method for identifying a compound which is effective in treating myeloma bone disease. The method comprises subjecting the compound to an assay to determine its ability to inhibit NF-kB activity or production, or its ability to inhibit proteasomal enzyme activity or production. The compounds reduce myeloma tumour volume, delay onset of limb paralysis, decrease the viability of myeloma cells and reduce the volume of tumour marker, Ibg2b. The compounds are useful for treating multiple myeloma such as osteopenia, osteolytic lesions, osteopetrosis, bone fracture and osteolytic bone disease, and myeloma bone disease
XX
SQ Sequence 4 AA;
Query Match 71.4%; Score 15; DB 3; Length 4;
Best Local Similarity 75.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GPSL 4
Db 1 GPFL 4

RESULT 12
ABG70996

ID XX
AC XX
XX ABG70996;
XX
DT 13-DEC-2002 (first entry)
DE
XX
XX Synthetic proteasome inhibitor peptide ALLM #1.
KW Hair growth; NF-kappaB; proteasomal; alopecia; human; mammal;
KW hair growth; animal; bone growth; degenerative bone disease; fracture;
KW dental problem; proteasome inhibitor; necrosis factor; ALLM.
XX
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "Modified by carbobenzoyl"
FT Modified-site 4 /note= "Modified by carbonyl group"
FT
XX US2002103127-A1.
XX
XX 01-AUG-2002.
XX
XX 15-JAN-2002; 2002US-00050425.
XX
XX 10-JUL-1998; 98US-00113947.
XX 27-JUL-1999; 99US-00361775.
XX (MUNDY) MUNDY G R.
XX (GARR) GARRETT I R.
XX (ROSS) ROSSINI G.
XX
XX Mundy GR, Garrett IR, Rossini G;
XX
XX WPI; 2002-739890/80.
XX
XX Stimulating hair growth, useful e.g. for treating alopecia, by administering compound that inhibits NF-kappaB or proteasomal enzymes.
XX
XX Disclosure; Page 5; 9pp; English.
XX
XX The present invention relates to a new method of stimulating hair growth in mammals by administering a compound that inhibits either activity of NF-kappaB, proteasomal activity or production of NF-kappaB or proteasomal proteins. The compounds of the invention are used to treat alopecia in humans and to increase hair growth in other animals. The invention can also be used to stimulate growth of bone e.g. for treating degenerative bone diseases, fractures and dental problems. The present amino acid sequence represents a peptide that was used in the invention to inhibit proteasomal activity
XX
SQ Sequence 4 AA;
Query Match 71.4%; Score 15; DB 5; Length 4;
Best Local Similarity 75.0%; Pred. No. 1.4e+06;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GPSL 4
Db 1 GPFL 4

RESULT 13
AAU76938
ID AAU76938 standard; peptide; 4 AA.
XX
XX AAU76938;
XX
XX
DT 05-JUN-2002 (first entry)
XX
XX NF-kappaB proteasome inhibitor.
XX

KW Proteasome inhibitor; Bone formation; hair growth; arthritis;
 KW Nuclear transcription factor beta inhibitor; osteopathic;
 KW antiinflammatory; vulnary; osteogenic; osteoporosis; bone fracture;
 KW hyperparathyroidism; periodontal disease; metastatic bone disease;
 KW osteolytic bone disease; male pattern baldness; alopecia areata.
 XX OS Synthetic.
 XX Key Location/Qualifiers
 FT Modified-site 1
 FT Modified-site 4 /note= "N-Carbobenzoyl"
 FT Modified-site 4 /note= "CHO modified"
 XX WO200128579-A2.
 XX 26-APR-2001.
 XX 20-OCT-2000; 2000WO-US041360.
 XX 20-OCT-1999; 99US-00421545.
 PR 25-APR-2000; 2000US-00558973.
 XX (OSTE-) OSTEOSCREEN INC.
 XX Mundy GR, Garrett RI, Rossini G;
 XX WPI; 2002-256022/30.
 XX Enhancing bone formation, treating pathological dental condition and
 FT degenerative joint condition e.g. osteoporosis involves use of a compound
 PT that inhibits nuclear transcription factor beta or proteasome activity.
 XX Disclosure; Page 26; 57pp; English.
 XX This invention relates to a novel method for enhancing bone formation.
 CC treating a pathological dental condition or treating degenerative joint
 CC conditions in a vertebrate animal. The method involves administration of
 CC a compound that inhibits the activity of a nuclear transcription factor
 CC beta (NF-kbeta), inhibits proteasomal activity or inhibits production of
 CC proteasomes. The invention may be used to treat pathological dental
 CC conditions, to treat degenerative joint conditions in a vertebrate animal
 CC such as osteoporosis, bone fracture or deficiency, primary or secondary
 CC hyperparathyroidism, periodontal disease or defect, metastatic bone
 CC disease, osteolytic bone disease, post-plastic surgery, post-prosthetic
 CC joint surgery and post-dental implantation; and for stimulating hair
 CC growth in a mammalian subject. The disorders of hair growth include male
 CC pattern baldness, alopecia areata, alopecia induced by cancer
 CC chemotherapy and hair thinning associated with ageing. The bone defects
 CC include elevation of peak bone mass in pre-menopausal women, growth
 CC deficiencies including age-related, post-menopausal, glucocorticoid
 CC induced osteoporosis and disease osteoporosis, arthritis, repair of
 CC congenital and trauma-induced resection of bone, for limiting or treating
 CC cartilage defects or disorders, and in wound healing and tissue repair.
 CC The administration of the compounds leads to increased bone growth and
 CC formation and stimulation of hair follicles. The compound does not
 CC inhibit the isoprenoid pathway. The present sequence represents an NF-
 CC kappaB or proteasome inhibitor which may be used in the method of the
 XX invention
 XX Sequence 4 AA;
 Query Match 71.4%; Score 15; DB 5; Length 4;
 Best Local Similarity 75.0%; Pred. No. 1.4e-06;
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GPSL 4
 Db 1 GPPL 4

RESULT 14
 ABG31209

ID ABG31209 standard; peptide; 4 AA.
 XX ABG31209;
 XX 05-NOV-2002 (first entry)
 XX Rat delta PKC first variable region, delta VI-1, modified peptide, #26.
 XX Rat; delta protein kinase C; deltaPKC; VI domain; vasotropic;
 KW cerebroprotective; deltaVI-1; deltaVI-2; RACK; pseudo-delta RACK;
 KW pseudo-delta receptor for activated C-kinase; deltaVI-5; PKC;
 KW protein kinase C; signal transduction; cell growth; gene expression;
 KW ion channel activity; translocation; hypoxia; stroke; ischaemic damage;
 KW creatine kinase; antagonist; agonist; mutant; mutein.
 XX Rattus norvegicus.
 OS Synthetic.
 XX Key Location/Qualifiers
 FT Misc-difference 2 /note= "Wild-type Gly substituted by Pro"
 FT WO200257413-A2.
 XX 25-JUL-2002.
 XX 09-NOV-2001; 2001WO-US047556.
 XX 18-JAN-2001; 2001US-0262060P.
 XX (STRD) UNIV LELAND STANFORD JUNIOR.
 XX Mochly-Rosen D;
 XX WPI; 2002-599715/64.
 XX New delta protein kinase C peptide for reducing or enhancing damage to
 FT cells or tissues exposed to ischemic or hypoxic event caused by stroke,
 FT or for protecting tissue from damage due to ischemia.
 XX Claim 5; Page 22; 65pp; English.
 XX The invention discloses peptides comprising deltaVI-1, deltaVI-2, pseudo-
 CC delta receptors for activated C-kinase (RACK), deltaVI-5 or their
 CC derivatives or fragments. Protein Kinase C (PKC) is a key enzyme in
 CC signal transduction involved in a variety of cellular functions including
 CC cell growth, regulation of gene expression and ion channel activity. The
 CC localisation of different PKC isozymes to different areas of the cell in
 CC turn appears due to binding of the activated isozymes to the specific
 CC anchoring molecules (RACKs). Peptides that mimic either the PKC-binding
 CC site on RACKs or the RACK-binding site on PKC are isozyme specific
 CC translocation inhibitors of PKC. The disclosed peptides are useful in
 CC activating or inhibiting translocation or function of deltaPKC. The
 CC deltaPKC agonists or antagonists are useful in reducing, enhancing or
 CC protecting against damage to cells or tissues due to ischaemic or hypoxic
 CC event caused by stroke. Acute administration of the peptides, conjugated
 CC to a carrier peptide or a Tat-derived peptide, protected hearts against
 CC ischaemic damage as shown by decreased release of creatine kinase. The
 CC data indicate that in an intact heart, inhibition of deltaPKC conferred
 CC greater than 50% protection against ischaemic damage. The peptides in
 CC ABG31159-ABG31219 are the deltaVI-1, deltaVI-2, pseudo-delta RACK,
 CC deltaVI-5 (or their derivatives or fragments) agonists or antagonists to
 CC the rat delta protein kinase C (PKC)
 XX Sequence 4 AA;
 Query Match 71.4%; Score 15; DB 5; Length 4;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 PSL 4
 Db 2 PSL 4

RESULT 15
 AAE29577
 ID AAE29577 standard; peptide; 4 AA.
 XX
 AC AAE29577;
 XX
 DT 27-JAN-2003 (first entry)
 XX
 DE Metallopeptide N-terminal peptide #8 used for Prion disease treatment.
 XX
 KW Metallopeptide; nontropic; amyloid beta-protein; Alzheimer's disease; AD;
 KW Prion's disease; oxytocin; angiotensin; vasopressin; neuroprotective;
 KW therapy.
 XX
 OS Unidentified.
 XX
 PN WO200264734-A2.
 XX
 PD 22-AUG-2002.
 XX
 PF 19-DEC-2001; 2001WO-US050075.
 XX
 PR 19-DEC-2000; 2000US-0256842P.
 PR 11-JUL-2001; 2001US-0304835P.
 PR 04-OCT-2001; 2001US-0327835P.
 XX
 PA (PALA-) PALATIN TECHNOLOGIES INC.
 XX
 PI Sharma SD, Shi Y;
 XX
 DR WPI; 2002-740699/80.
 XX
 FT Determining secondary structure binding to desired targets within parent
 FT polypeptides that bind to targets, by constructing and complexing
 FT peptides to metal ions to form metallopeptides and screening the
 FT metallopeptides.
 XX
 PS Example 3; Page 142; 165pp; English.
 XX
 CC The invention relates to a method for identification and determination of
 CC target-specific folding sites in peptides and proteins. The invention
 CC also relates to a method for determining a secondary structure binding to
 CC desired targets within parent polypeptides that bind to targets, by
 CC constructing and complexing peptides to metal ions to form
 CC metallopeptides and screening the metallopeptides. The method is useful
 CC for determining secondary structure binding to desired target within
 CC parent polypeptide with primary structure that binds to the target, where
 CC the target of interest is a receptor, antibody, toxin, enzyme, hormone,
 CC nucleic acid, intracellular protein domain of biological relevance or
 CC extracellular protein domain of biological relevance. A library of
 CC amyloid beta-protein related peptides is useful for the treatment of
 CC Alzheimer's disease (AD). A library of peptides targeting vasopressin,
 CC oxytocin or angiotensin receptor is useful for treating prion's disease.
 CC The present sequence is a metallopeptide N-terminal peptide used for
 CC Prion disease treatment. This peptide is used to illustrate the method of
 CC the invention
 XX
 SQ Sequence 4 AA;

Query Match 71.4%; Score 15; DB 5; Length 4;
 Best Local Similarity 50.0%; Pred. No. 1.4e+06;
 Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GPSL 4
 ||:
 Db 1 GPAV 4

Search completed: March 10, 2004, 15:06:37
 Job time : 54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 15:07:56 ; Search time 33 Seconds
(without alignments)
25.594 Million cell updates/sec

Title: US-09-848-834A-5

Perfect score: 21

Sequence: 1 GPSL 4

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 89742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 5686

Minimum DB seq length: 0

Maximum DB seq length: 4

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	4	9	US-09-848-834A-5
2	21	100.0	4	14	US-10-285-976-224
3	15	71.4	4	13	US-10-007-761-61
4	13	61.9	2	10	US-09-836-433-15
5	13	61.9	3	9	US-09-967-003-1
6	13	61.9	3	10	US-09-836-433-16
7	13	61.9	3	10	US-09-755-630A-277
8	13	61.9	3	14	US-10-283-838-1
9	13	61.9	4	9	US-09-040-518-9
10	13	61.9	4	9	US-09-804-733A-24
11	13	61.9	4	9	US-09-258-650-4
12	13	61.9	4	9	US-09-765-614B-30
13	13	61.9	4	9	US-09-925-715-26
14	13	61.9	4	9	US-09-269-439-10
15	13	61.9	4	9	US-09-269-439-11

Sequence 71, Appl
Sequence 16, Appl
Sequence 3, Appl
Sequence 12, Appl
Sequence 26, Appl
Sequence 27, Appl
Sequence 17, Appl
Sequence 71, Appl
Sequence 158, Appl
Sequence 2, Appl
Sequence 18, Appl
Sequence 11, Appl
Sequence 3, Appl
Sequence 26, Appl
Sequence 27, Appl
Sequence 10, Appl
Sequence 6, Appl
Sequence 3, Appl
Sequence 4, Appl
Sequence 5, Appl
Sequence 9, Appl
Sequence 11, Appl
Sequence 19, Appl
Sequence 17, Appl
Sequence 84, Appl
Sequence 1, Appl
Sequence 26, Appl
Sequence 27, Appl
Sequence 45, Appl

4 9 US-09-908-322-71
4 9 US-09-779-054-16
4 9 US-09-916-940-3
4 9 US-09-947-387-12
4 9 US-09-972-772-26
4 9 US-09-972-772-27
4 10 US-09-836-433-17
4 10 US-09-783-931-71
4 10 US-09-852-910-158
4 10 US-09-931-009A-2
4 10 US-09-933-025-18
4 10 US-09-893-525-11
4 11 US-09-807-742-3
4 13 US-10-001-945-26
4 13 US-10-001-945-27
4 14 US-10-057-467-10
4 14 US-10-046-801-6
4 14 US-10-036-111-3
4 14 US-10-036-111-4
4 14 US-10-036-111-5
4 14 US-10-036-111-9
4 14 US-10-036-111-10
4 14 US-10-286-516A-11
4 14 US-10-158-742A-19
4 14 US-10-117-931-17
4 14 US-10-128-590-84
4 14 US-10-234-319A-1
4 14 US-10-138-935-26
4 14 US-10-138-935-27
4 14 US-10-001-073-45

16 13 61.9
17 13 61.9
18 13 61.9
19 13 61.9
20 13 61.9
21 13 61.9
22 13 61.9
23 13 61.9
24 13 61.9
25 13 61.9
26 13 61.9
27 13 61.9
28 13 61.9
29 13 61.9
30 13 61.9
31 13 61.9
32 13 61.9
33 13 61.9
34 13 61.9
35 13 61.9
36 13 61.9
37 13 61.9
38 13 61.9
39 13 61.9
40 13 61.9
41 13 61.9
42 13 61.9
43 13 61.9
44 13 61.9
45 13 61.9

ALIGNMENTS

RESULT 1
US-09-848-834A-5
; Sequence 5, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aptton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 5
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
US-09-848-834A-5

Query Match 100.0%; Score 21; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPSL 4
Db 1 GPSL 4

RESULT 2
US-10-285-976-224
; Sequence 224, Application US/10285976
; Publication No. US20030165500A1
; GENERAL INFORMATION:
; APPLICANT: Rhee, Chae-Seo
; APPLICANT: Malini, Sen

; APPLICANT: Wu, Christina
 ; APPLICANT: Leoni, Lorenzo M.
 ; APPLICANT: Cort, Maribat
 ; APPLICANT: Carlson, Dennis A.
 ; APPLICANT: The Regents of the University of California
 ; TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy
 ; TITLE OF INVENTION: in Head and Neck Squamous Cell Carcinomas
 ; FILE REFERENCE: 023070-130320US
 ; CURRENT APPLICATION NUMBER: US/10/285,976
 ; CURRENT FILING DATE: 2002-11-01
 ; PRIOR APPLICATION NUMBER: US 60/287,995
 ; PRIOR FILING DATE: 2001-05-01
 ; PRIOR APPLICATION NUMBER: WO PCT/US02/13802
 ; PRIOR FILING DATE: 2002-05-01
 ; NUMBER OF SEQ ID NOS: 232
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 224
 ; LENGTH: 4
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:short linker
 ; OTHER INFORMATION: sequence
 ; US-10-285-976-224

Query Match 100.0%; Score 21; DB 14; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.1e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GPSL 4
 ||||
 Db 1 GPSL 4

RESULT 3
 US-10-007-761-61
 ; Sequence 61, Application US/10007761
 ; Publication No. US20030150984A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mochly-Rosen, Daria
 ; TITLE OF INVENTION: Peptides for Activation and Inhibition
 ; TITLE OF INVENTION: of delta-PKC
 ; FILE REFERENCE: 58600-8208.US00
 ; CURRENT APPLICATION NUMBER: US/10/007,761
 ; CURRENT FILING DATE: 2001-11-09
 ; PRIOR APPLICATION NUMBER: US 60/262,060
 ; PRIOR FILING DATE: 2001-01-18
 ; NUMBER OF SEQ ID NOS: 72
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 61
 ; LENGTH: 4
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: modified fragment of delta V1-1 peptide
 ; US-10-007-761-61

Query Match 71.4%; Score 15; DB 13; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.1e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PSL 4
 ||||
 Db 2 PSL 4

RESULT 4
 US-09-836-433-15
 ; Sequence 15, Application US/09836433
 ; Publication No. US20030049797A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yuki, Yoshikazu
 ; APPLICANT: Udagawa, Shigezo

; TITLE OF INVENTION: HYBRID PROTEINS FOR AUTOIMMUNE DISEASE
 ; FILE REFERENCE: No. US20030049797A1 Assigned
 ; CURRENT APPLICATION NUMBER: US/09/836,433
 ; CURRENT FILING DATE: 2001-04-16
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 15
 ; LENGTH: 2
 ; TYPE: PRT
 ; ORGANISM: synthetic construct
 ; US-09-836-433-15

Query Match 61.9%; Score 13; DB 10; Length 2;
 Best Local Similarity 100.0%; Pred. No. 7.1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GP 2
 ||
 Db 1 GP 2

RESULT 5
 US-09-967-003-1
 ; Sequence 1, Application US/09967003
 ; Patent No. US20020107202A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haddox, Jeffrey
 ; APPLICANT: Pfister, Robert
 ; APPLICANT: Bialock, James
 ; APPLICANT: Matteo, Villain
 ; TITLE OF INVENTION: SYNTHETIC COMPLEMENTARY PEPTIDES AND OPHTHALMOLOGIC
 ; FILE REFERENCE: 92750/57
 ; CURRENT APPLICATION NUMBER: US/09/967,003
 ; CURRENT FILING DATE: 2001-09-28
 ; PRIOR APPLICATION NUMBER: US/09/521,365
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: US 60/123,409
 ; PRIOR FILING DATE: 1999-03-08
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 3
 ; TYPE: PRT
 ; ORGANISM: unknown organism
 ; FEATURE:
 ; OTHER INFORMATION: unknown organism: amino acid sequence of neutrophilic
 ; OTHER INFORMATION: chemoattractant
 ; OTHER INFORMATION: released during direct alkaline hydrolysis of corneal proteins;
 ; OTHER INFORMATION: polymorphonuclear leukocyte invasion into alkali-injured cornea
 ; US-09-967-003-1

Query Match 61.9%; Score 13; DB 9; Length 3;
 Best Local Similarity 100.0%; Pred. No. 7.1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GP 2
 ||
 Db 2 GP 3

RESULT 6
 US-09-836-433-16
 ; Sequence 16, Application US/09836433
 ; Publication No. US20030049797A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yuki, Yoshikazu
 ; APPLICANT: Udagawa, Shigezo
 ; TITLE OF INVENTION: HYBRID PROTEINS FOR AUTOIMMUNE DISEASE
 ; FILE REFERENCE: No. US20030049797A1 Assigned
 ; CURRENT APPLICATION NUMBER: US/09/836,433
 ; CURRENT FILING DATE: 2001-04-16
 ; NUMBER OF SEQ ID NOS: 49

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 3
; TYPE: PRT
; ORGANISM: synthetic construct
US-09-836-433-16

Query Match
Best Local Similarity 61.9%; Score 13; DB 10; Length 3;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GP 2
DB 1 GP 2

RESULT 7
US-09-755-630A-277
; Sequence 277, Application US/09755630A
; Publication No. US20030194399A1
; GENERAL INFORMATION:
; APPLICANT: ALIBAI, MURTAZA F.
; APPLICANT: ASTWOOD, JAMES D.
; APPLICANT: SAMPSON, HUGH A.
; APPLICANT: MCWHERTER, CHARLES A.
; TITLE OF INVENTION: PREPARATION OF DEALLERGENIZED PROTEINS AND PERMUTEINS
; FILE REFERENCE: 11899.0217.NPUS00 (NOB0217)
; CURRENT APPLICATION NUMBER: US/09/755,630A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,669
; PRIOR FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 293
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 277
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-09-755-630A-277

Query Match
Best Local Similarity 61.9%; Score 13; DB 10; Length 3;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GP 2
DB 1 GP 2

RESULT 8
US-10-283-838-1
; Sequence 1, Application US/10283838
; Publication No. US2003092894A1
; GENERAL INFORMATION:
; APPLICANT: Per Antonsson, Per Bjork, Mikael Dohlsten,
; Johan Hansson, Terje Kalland, Lars
; Abrahamson and Goran Forsberg
; TITLE OF INVENTION: MODIFIED/CHIMERIC SUPERANTIGENS
; AND THEIR USE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pravei, Hewitt, Kimball & Krieger
; STREET: 1177 West Loop South, 10th Floor
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77027-9095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
US-10-283-838-1

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 3
; TYPE: PRT
; ORGANISM: synthetic construct
US-09-836-433-16

Query Match
Best Local Similarity 61.9%; Score 13; DB 10; Length 3;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GP 2
DB 1 GP 2

RESULT 7
US-09-755-630A-277
; Sequence 277, Application US/09755630A
; Publication No. US20030194399A1
; GENERAL INFORMATION:
; APPLICANT: ALIBAI, MURTAZA F.
; APPLICANT: ASTWOOD, JAMES D.
; APPLICANT: SAMPSON, HUGH A.
; APPLICANT: MCWHERTER, CHARLES A.
; TITLE OF INVENTION: PREPARATION OF DEALLERGENIZED PROTEINS AND PERMUTEINS
; FILE REFERENCE: 11899.0217.NPUS00 (NOB0217)
; CURRENT APPLICATION NUMBER: US/09/755,630A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,669
; PRIOR FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 293
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 277
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-09-755-630A-277

Query Match
Best Local Similarity 61.9%; Score 13; DB 10; Length 3;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GP 2
DB 1 GP 2

RESULT 8
US-10-283-838-1
; Sequence 1, Application US/10283838
; Publication No. US2003092894A1
; GENERAL INFORMATION:
; APPLICANT: Per Antonsson, Per Bjork, Mikael Dohlsten,
; Johan Hansson, Terje Kalland, Lars
; Abrahamson and Goran Forsberg
; TITLE OF INVENTION: MODIFIED/CHIMERIC SUPERANTIGENS
; AND THEIR USE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pravei, Hewitt, Kimball & Krieger
; STREET: 1177 West Loop South, 10th Floor
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77027-9095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
US-10-283-838-1

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 3
; TYPE: PRT
; ORGANISM: synthetic construct
US-09-836-433-16

Query Match
Best Local Similarity 61.9%; Score 13; DB 10; Length 3;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GP 2
DB 1 GP 2

RESULT 7
US-09-755-630A-277
; Sequence 277, Application US/09755630A
; Publication No. US20030194399A1
; GENERAL INFORMATION:
; APPLICANT: ALIBAI, MURTAZA F.
; APPLICANT: ASTWOOD, JAMES D.
; APPLICANT: SAMPSON, HUGH A.
; APPLICANT: MCWHERTER, CHARLES A.
; TITLE OF INVENTION: PREPARATION OF DEALLERGENIZED PROTEINS AND PERMUTEINS
; FILE REFERENCE: 11899.0217.NPUS00 (NOB0217)
; CURRENT APPLICATION NUMBER: US/09/755,630A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,669
; PRIOR FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 293
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 277
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-09-755-630A-277

Query Match
Best Local Similarity 61.9%; Score 13; DB 10; Length 3;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GP 2
DB 1 GP 2

RESULT 8
US-10-283-838-1
; Sequence 1, Application US/10283838
; Publication No. US2003092894A1
; GENERAL INFORMATION:
; APPLICANT: Per Antonsson, Per Bjork, Mikael Dohlsten,
; Johan Hansson, Terje Kalland, Lars
; Abrahamson and Goran Forsberg
; TITLE OF INVENTION: MODIFIED/CHIMERIC SUPERANTIGENS
; AND THEIR USE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pravei, Hewitt, Kimball & Krieger
; STREET: 1177 West Loop South, 10th Floor
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77027-9095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
US-10-283-838-1

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 3
; TYPE: PRT
; ORGANISM: synthetic construct
US-09-836-433-16

Query Match
Best Local Similarity 61.9%; Score 13; DB 10; Length 3;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GP 2
DB 1 GP 2

RESULT 7
US-09-755-630A-277
; Sequence 277, Application US/09755630A
; Publication No. US20030194399A1
; GENERAL INFORMATION:
; APPLICANT: ALIBAI, MURTAZA F.
; APPLICANT: ASTWOOD, JAMES D.
; APPLICANT: SAMPSON, HUGH A.
; APPLICANT: MCWHERTER, CHARLES A.
; TITLE OF INVENTION: PREPARATION OF DEALLERGENIZED PROTEINS AND PERMUTEINS
; FILE REFERENCE: 11899.0217.NPUS00 (NOB0217)
; CURRENT APPLICATION NUMBER: US/09/755,630A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,669
; PRIOR FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 293
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 277
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-09-755-630A-277

Query Match
Best Local Similarity 61.9%; Score 13; DB 10; Length 3;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GP 2
DB 1 GP 2

RESULT 8
US-10-283-838-1
; Sequence 1, Application US/10283838
; Publication No. US2003092894A1
; GENERAL INFORMATION:
; APPLICANT: Per Antonsson, Per Bjork, Mikael Dohlsten,
; Johan Hansson, Terje Kalland, Lars
; Abrahamson and Goran Forsberg
; TITLE OF INVENTION: MODIFIED/CHIMERIC SUPERANTIGENS
; AND THEIR USE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pravei, Hewitt, Kimball & Krieger
; STREET: 1177 West Loop South, 10th Floor
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77027-9095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
US-10-283-838-1

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 3
; TYPE: PRT
; ORGANISM: synthetic construct
US-09-836-433-16

Query Match
Best Local Similarity 61.9%; Score 13; DB 10; Length 3;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GP 2
DB 1 GP 2

RESULT 7
US-09-755-630A-277
; Sequence 277, Application US/09755630A
; Publication No. US20030194399A1
; GENERAL INFORMATION:
; APPLICANT: ALIBAI, MURTAZA F.
; APPLICANT: ASTWOOD, JAMES D.
; APPLICANT: SAMPSON, HUGH A.
; APPLICANT: MCWHERTER, CHARLES A.
; TITLE OF INVENTION: PREPARATION OF DEALLERGENIZED PROTEINS AND PERMUTEINS
; FILE REFERENCE: 11899.0217.NPUS00 (NOB0217)
; CURRENT APPLICATION NUMBER: US/09/755,630A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,669
; PRIOR FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 293
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 277
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-09-755-630A-277

Query Match
Best Local Similarity 61.9%; Score 13; DB 10; Length 3;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GP 2
DB 1 GP 2

RESULT 8
US-10-283-838-1
; Sequence 1, Application US/10283838
; Publication No. US2003092894A1
; GENERAL INFORMATION:
; APPLICANT: Per Antonsson, Per Bjork, Mikael Dohlsten,
; Johan Hansson, Terje Kalland, Lars
; Abrahamson and Goran Forsberg
; TITLE OF INVENTION: MODIFIED/CHIMERIC SUPERANTIGENS
; AND THEIR USE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pravei, Hewitt, Kimball & Krieger
; STREET: 1177 West Loop South, 10th Floor
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77027-9095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
US-10-283-838-1

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 3
; TYPE: PRT
; ORGANISM: synthetic construct
US-09-836-433-16
```

```

; CURRENT APPLICATION NUMBER: US/09/804,733A
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: US 60/188,990
; PRIOR FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; NAME/KEY: VARIANT
; LOCATION: (1)..(4)
; OTHER INFORMATION: Trypsin cleavage site
US-09-804-733A-24

Query Match          61.9%; Score 13; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GP 2
Db 2 GP 3

RESULT 11
US-09-256-650-4
; Sequence 4, Application US/09256650
; Patent No. US20020098578A1
; GENERAL INFORMATION:
; APPLICANT: Prockop, Darwin J.
; APPLICANT: Ala-Kokko, Leena
; APPLICANT: Fertala, Andrzej
; APPLICANT: Sieron, Aleksander
; APPLICANT: Kivirikko, Kari I.
; APPLICANT: Geddis, Amy
; TITLE OF INVENTION: Synthesis of Human Procollagens
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. US20020098578A1ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/256,650
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/211,820
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/780,899
; FILING DATE: 23-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

US-09-256-650-4
Query Match          61.9%; Score 13; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GP 2
Db 2 GP 3

RESULT 12
US-09-765-614B-30
; Sequence 30, Application US/09765614B
; Patent No. US20020102215A1
; GENERAL INFORMATION:
; APPLICANT: Nycomed Imaging AS
; TITLE OF INVENTION: Improvements in or relating to
; TITLE OF INVENTION: diagnostic/therapeutic
; TITLE OF INVENTION: agents
; FILE REFERENCE: REF/Klaveness/054
; CURRENT APPLICATION NUMBER: US/09/765,614B
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial
; OTHER INFORMATION: Sequence:Thrombus
; OTHER INFORMATION: binding peptide
US-09-765-614B-30

Query Match          61.9%; Score 13; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GP 2
Db 1 GP 2

RESULT 13
US-09-925-715-26
; Sequence 26, Application US/09925715
; Patent No. US20020102217A1
; GENERAL INFORMATION:
; APPLICANT: Nycomed Imaging AS
; TITLE OF INVENTION: Improvements in or relating to diagnostic/therapeutic
; FILE REFERENCE: REF/Klaveness/206
; CURRENT APPLICATION NUMBER: US/09/925,715
; CURRENT FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Thrombus
; OTHER INFORMATION: binding peptide
US-09-925-715-26

Query Match          61.9%; Score 13; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GP 2
Db 1 GP 2
```

RESULT 14

US-09-269-439-10
; Sequence 10, Application US/09269439
; Patent No. US20020107175A1
; GENERAL INFORMATION:
; APPLICANT: Wahren, John
; APPLICANT: Johansson, Bo-Lennart
; APPLICANT: Joinvall, Hans
; TITLE OF INVENTION: INSULIN C-PEPTIDES
; FILE REFERENCE: 08269/008001
; CURRENT APPLICATION NUMBER: US/09/269,439
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: PCT/GB97/02627
; EARLIER FILING DATE: 1997-09-26
; EARLIER APPLICATION NUMBER: SE96/03533-2
; EARLIER FILING DATE: 1996-09-27
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-269-439-10

Query Match 61.9%; Score 13; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GP 2
||
Db 3 GP 4

RESULT 15

US-09-269-439-11
; Sequence 11, Application US/09269439
; Patent No. US20020107175A1
; GENERAL INFORMATION:
; APPLICANT: Wahren, John
; APPLICANT: Johansson, Bo-Lennart
; APPLICANT: Joinvall, Hans
; TITLE OF INVENTION: INSULIN C-PEPTIDES
; FILE REFERENCE: 08269/008001
; CURRENT APPLICATION NUMBER: US/09/269,439
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: PCT/GB97/02627
; EARLIER FILING DATE: 1997-09-26
; EARLIER APPLICATION NUMBER: SE96/03533-2
; EARLIER FILING DATE: 1996-09-27
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-269-439-11

Query Match 61.9%; Score 13; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GP 2
||
Db 2 GP 3

Search completed: March 10, 2004, 15:13:11
Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 19.7451 Seconds
(without alignments)
133.345 Million cell updates/sec

Title: US-09-848-834A-20

Perfect score: 265

Sequence: 1 XHWSYGLRPGSSGSLDEKX.....NVVNSSGGPSLHWSYGLRFX 51

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCPUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	127.5	48.1	33	1 US-08-446-692-27
2	127.5	48.1	33	2 US-08-488-351A-27
3	99	37.4	412	1 US-08-313-288B-18
4	99	37.4	424	2 US-08-760-797A-3
5	99	37.4	424	3 US-08-932-929B-3
6	97.5	36.8	49	1 US-08-387-156-4
7	97.5	36.8	49	2 US-08-694-865-4
8	97.5	36.8	49	3 US-08-878-748-4
9	97.5	36.8	49	4 US-08-124-491-4
10	97.5	36.8	49	5 US-08-383-912-4
11	97.5	36.8	544	1 US-08-387-156-10
12	97.5	36.8	544	2 US-08-694-865-10
13	97.5	36.8	544	3 US-08-878-748-10
14	97.5	36.8	544	4 US-08-124-491-10
15	97.5	36.8	544	5 US-08-383-912-10
16	97.5	36.8	699	2 US-08-694-865-16
17	97.5	36.8	699	3 US-08-124-491-16
18	97.5	36.8	699	4 US-08-383-912-16
19	97.5	36.8	977	1 US-08-387-156-8
20	97.5	36.8	977	2 US-08-694-865-8
21	97.5	36.8	977	3 US-08-878-748-8
22	97.5	36.8	977	4 US-08-124-491-8
23	97.5	36.8	977	5 US-08-383-912-8
24	91	34.3	423	2 US-08-760-797A-1
25	91	34.3	424	3 US-08-932-929B-1
26	89	33.6	21	1 US-08-186-266-6
27	89	33.6	21	2 US-08-446-692-48

28	89	33.6	21	2	US-08-488-351A-48	Sequence 48, Appl
29	89	33.6	21	3	US-09-100-409A-54	Sequence 54, Appl
30	89	33.6	21	4	US-08-464-436-17	Sequence 17, Appl
31	89	33.6	21	4	US-08-788-822A-12	Sequence 12, Appl
32	89	33.6	21	4	US-08-197-484-97	Sequence 97, Appl
33	89	33.6	21	4	US-09-543-608A-39	Sequence 39, Appl
34	89	33.6	21	5	PCT-US95-02121-97	Sequence 20, Appl
35	89	33.6	21	5	PCT-US95-13841-20	Sequence 97, Appl
36	88.5	33.4	40	4	US-09-026-276-35	Sequence 35, Appl
37	88.5	33.4	40	4	US-09-964-201A-35	Sequence 35, Appl
38	88.5	33.4	41	4	US-09-026-276-34	Sequence 34, Appl
39	88.5	33.4	41	4	US-09-964-201A-34	Sequence 34, Appl
40	87.5	33.0	44	1	US-07-690-983D-45	Sequence 45, Appl
41	87.5	33.0	52	3	US-08-458-814-6	Sequence 6, Appl
42	87.5	33.0	55	3	US-08-458-814-7	Sequence 7, Appl
43	87.5	33.0	84	1	US-07-690-983D-47	Sequence 47, Appl
44	84.5	31.9	20	1	US-07-690-983D-40	Sequence 40, Appl
45	84.5	31.9	20	4	US-09-026-276-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1

US-08-446-692-27
; Sequence 27, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-27

Query Match 48.1%; Score 127.5; DB 1; Length 33;
Best Local Similarity 81.8%; Pred. No. 1.2e-09;
Matches 27; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

OY 18 EKXIAMEKASSVFNNVNSSGSPSLHWSYGLRP 50

Db 3 EKXIAMEKASSVFNNVNSSGSPSLHWSYGLRP 32

```
RESULT 2
US-08-488-351A-27
; Sequence 27, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488.351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-27
Query Match 48.1%; Score 127.5; DB 2; Length 33;
Best Local Similarity 81.8%; Pred. No. 1.2e-09;
Matches 27; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 18 EKXIAKMEKASSVFNVNSSGCSPLHMSYGLRP 50
DB 3 EKXIAKMEKASSVFNVNSSGCSPLHMSYGLRP 32

RESULT 3
US-08-313-288B-18
; Sequence 18, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, P-SPONDIN
; NUMBER OF SEQUENCES: 20
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313.288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-18
Query Match 37.4%; Score 99; DB 1; Length 412;
Best Local Similarity 56.1%; Pred. No. 0.00012;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKXIAKMEKASSVFNVNSSSG 39
DB 361 IKFGSANKPKDLDYENIEKKICWKCSSVFNVNSSIG 401

RESULT 4
US-08-760-797A-3
; Sequence 3, Application US/08760797A
; Patent No. 5928902
; GENERAL INFORMATION:
; APPLICANT: De Wilde, Michel
; APPLICANT: Cohen, Joseph
; TITLE OF INVENTION: Hybrid Protein Between CS
; TITLE OF INVENTION: from Plasmodium and HBsAg
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,797A
; FILING DATE: 04-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Kirk
```

REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: B45015-1C2
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-760-797A-3

Query Match 37.4%; Score 99; DB 2; Length 424;
Best Local Similarity 56.1%; Pred. No. 0.00013;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPGSSGPSLD-----EKKIAMEKASSVFNVNSSSG 39
Db 152 IKPGSANKPKDLDYANDIEKKICKMEKCSVFNVNSSIG 192

RESULT 5

US-08-932-929B-3
Sequence 3, Application US/08932929B
Patent No. 6169171

GENERAL INFORMATION:
APPLICANT: De Wilde, Michel
APPLICANT: Cohen, Joseph
TITLE OF INVENTION: Hybrid Protein Between CS
TITLE OF INVENTION: from Plasmodium and HBSAG
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
FILING DATE: US/08/932.929B

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/760,797
FILING DATE: 04-DEC-1996
APPLICATION NUMBER: 08/442,612
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Kirk
REGISTRATION NUMBER: 33,833
REFERENCE/DOCKET NUMBER: B45015-1FWC2
TELEPHONE: 610-270-5096
TELEFAX: 610-270-5090
TELEX:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-932-929B-3

Query Match 37.4%; Score 99; DB 3; Length 424;
Best Local Similarity 56.1%; Pred. No. 0.00013;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPGSSGPSLD-----EKKIAMEKASSVFNVNSSSG 39
Db 152 IKPGSANKPKDLDYANDIEKKICKMEKCSVFNVNSSIG 192

RESULT 6

US-08-387-156-4
Sequence 4, Application US/08387156
Patent No. 5723129

GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HUW P.A.
TITLE OF INVENTION: GRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/387,156
FILING DATE: 10-FEB-1995

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991

ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-387-156-4

Query Match 36.8%; Score 97.5; DB 1; Length 49;
Best Local Similarity 42.9%; Pred. No. 1.4e-05;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

QY 2 HWSYGLRPGSSGSLDDEKKIAMEKASSVFNVNSSSGPSLHWSYGLRP 50
Db 2 HWSYGLRPGS-----GSQDWSYGLRPGSSQHSYGLRP 35

RESULT 7

US-08-694-865-4
Sequence 4, Application US/08694865
Patent No. 5837468

GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANNS, JOHN G.
TITLE OF INVENTION: GRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP


```

; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,865
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-694-865-4

Query Match 36.8%; Score 97.5; DB 2; Length 49;
Best Local Similarity 42.9%; Pred. No. 1.4e-05;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

QY 2 HWSYGLRPGSGSLDEKKIAKMEKASSVFNVNSSGSPSLHWSYGLRP 50
Db 2 HWSYGLRPGS-----GSQDWSYGLRPGSGSQHWSYGLRP 35

RESULT 8
US-08-878-748-4
; Sequence 4, Application US/08878748
; Patent No. 5969126
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: GRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,748
; FILING DATE: 19-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-124-491-4

Query Match 36.8%; Score 97.5; DB 2; Length 49;
Best Local Similarity 42.9%; Pred. No. 1.4e-05;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

QY 2 HWSYGLRPGSGSLDEKKIAKMEKASSVFNVNSSGSPSLHWSYGLRP 50
Db 2 HWSYGLRPGS-----GSQDWSYGLRPGSGSQHWSYGLRP 35

RESULT 9
US-09-124-491-4
; Sequence 4, Application US/09124491
; Patent No. 6022960
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNIS, JOHN G.
; TITLE OF INVENTION: GRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,491
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE: 09-AUG-1996
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-124-491-4
```

Query Match 36.8%; Score 97.5; DB 3; Length 49;
Best Local Similarity 42.9%; Pred. No. 1.4e-05;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

QY 2 HWSYGLRPGSGPSLDEKKIAKMEKASSVFNVNSSSGPSLHWSYGLRP 50
|||||
Db 2 HWSYGLRPGS-----GSQDWSYGLRPGSGSCHWSYGLRP 35

RESULT 10
US-09-383-912-4
; Sequence 4, Application US/09383912
; Patent No. 6521746
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNS, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-383-912-4

Query Match 36.8%; Score 97.5; DB 4; Length 49;
Best Local Similarity 42.9%; Pred. No. 1.4e-05;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

QY 2 HWSYGLRPGSGPSLDEKKIAKMEKASSVFNVNSSSGPSLHWSYGLRP 50
|||||
Db 2 HWSYGLRPGS-----GSQDWSYGLRPGSGSCHWSYGLRP 35

RESULT 11
US-08-387-156-10
; Sequence 10, Application US/08387156
; Patent No. 5723129
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HW P.A.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:

Query Match 36.8%; Score 97.5; DB 1; Length 544;
Best Local Similarity 42.9%; Pred. No. 0.00027;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

QY 2 HWSYGLRPGSGPSLDEKKIAKMEKASSVFNVNSSSGPSLHWSYGLRP 50
|||||
Db 495 HWSYGLRPGS-----GSQDWSYGLRPGSGSCHWSYGLRP 528

RESULT 12
US-08-694-865-10
; Sequence 10, Application US/08694865
; Patent No. 5837268
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNS, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,865
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-387-156-10

REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-865-10

Query Match 36.8%; Score 97.5; DB 2; Length 544;
Best Local Similarity 42.9%; Pred. No. 0.00027;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

QY 2 HWSYGLRPGSGPSLDEKIAKMEKASSVFNVNSSGSPSLHWSYGLRP 50
DB 495 HWSYGLRPGS-----GSQDWSYGLRPGSSQHWYGLRP 528

RESULT 13

US-08-878-748-10
Sequence 10, Application US/08878748
Patent No. 5969126

GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HUW P.A.
TITLE OF INVENTION: GARH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,748
FILING DATE: 19-JUN-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991

ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231

INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-878-748-10

Query Match 36.8%; Score 97.5; DB 2; Length 544;
Best Local Similarity 42.9%; Pred. No. 0.00027;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

QY 2 HWSYGLRPGSGPSLDEKIAKMEKASSVFNVNSSGSPSLHWSYGLRP 50
DB 495 HWSYGLRPGS-----GSQDWSYGLRPGSSQHWYGLRP 528

RESULT 14

US-09-124-491-10
Sequence 10, Application US/09124491
Patent No. 6022960

GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW A.
APPLICANT: MANNS, JOHN G.
TITLE OF INVENTION: GARH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,491
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE: 09-AUG-1996
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-491-10

Query Match 36.8%; Score 97.5; DB 3; Length 544;
Best Local Similarity 42.9%; Pred. No. 0.00027;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

QY 2 HWSYGLRPGSGPSLDEKIAKMEKASSVFNVNSSGSPSLHWSYGLRP 50
DB 495 HWSYGLRPGS-----GSQDWSYGLRPGSSQHWYGLRP 528

RESULT 15

US-09-383-912-10
Sequence 10, Application US/09383912
Patent No. 6521746

GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW A.
APPLICANT: MANNS, JOHN G.

Wed Mar 10 10:34:31 2004

TITLE OF INVENTION: GRRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/383,912
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-383-912-10

Query Match 36.8%; Score 97.5; DB 4; Length 544;
Best Local Similarity 42.9%; Pred. No. 0.00027;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

Cy 2 HWSYGLRPGSGPSLDEKKIAKMEKASVFNVNNSGSPSLHWSYGLRP 50
|||||
Db 495 HWSYGLRPGS-----GSQDWSYGLRPGSSQHSYGLRP 528

Search completed: March 10, 2004, 09:28:57
Job time : 20.7451 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:53 ; Search time 9.32685 Seconds
(without alignments)
284.724 Million cell updates/sec

Title: US-09-848-834A-20
Perfect score: 265
Sequence: 1 XHWSYGLRPGSSGFLDEKK.....NVNSSSGPSLHWSYGLRFX 51

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	39.6	388	1 CSP_PLARE	P26694 plasmodium
2	99	37.4	397	1 CSP_PLARO	P19587 plasmodium
3	99	37.4	412	1 CSP_PLAFA	P02893 plasmodium
4	99	37.4	424	1 CSP_PLAFT	P13814 plasmodium
5	98	37.0	442	1 CSP_PLAFW	P08307 plasmodium
6	66	24.9	339	1 CSP_PLABE	P06915 plasmodium
7	66	24.9	347	1 CSP_PLABA	P23093 plasmodium
8	63	23.8	367	1 CSP_PLAYO	P06914 plasmodium
9	62.5	23.6	82	1 GON3_SALTR	P45653 salmo trutt
10	62	23.4	63	1 GON1_MESAU	O09183 mesocricetu
11	61.5	23.2	61	1 GON1_SHEEP	Q28588 ovis aries
12	61.5	23.2	90	1 GON3_SPAAU	P51923 sparus aura
13	60.5	22.8	90	1 GON3_DICLA	Q91A09 dicentrarch
14	60	22.6	94	1 GON1_HAPBU	P51918 haplochromi
15	59.5	22.5	721	1 THIC_SHEON	O8ee47 shewanella
16	59	22.3	393	1 CSP_PLABR	P14593 plasmodium
17	59	22.3	429	1 CSP_PLAWA	P13815 plasmodium
18	58.5	22.1	74	1 GON3_ONCTS	Q92097 oncorhynch
19	58.5	22.1	82	1 GON3_SALSA	P35629 salmo salar
20	58.5	22.1	624	1 PLB2_SCHPO	O13857 schizosacch
21	58	21.9	67	1 GON1_MACWU	P55247 macaca mula
22	58	21.9	89	1 GON1_XENLA	P45656 xenopus lae
23	58	21.9	90	1 GON1_MOUSE	P13582 mus musculus
24	58	21.9	90	1 GON1_RANCA	Q90583 rana catesb
25	58	21.9	91	1 GON1_PIG	P49921 sus scrofa
26	58	21.9	92	1 GON1_HUMAN	P01148 homo sapien
27	58	21.9	92	1 GON1_RAT	P07490 rattus norv
28	58	21.9	92	1 GON1_TURGB	Q95335 tupai
29	58	21.9	74	1 TRME_BORBU	P53364 borrella bu
30	57.5	21.7	464	1 GON3_ONCMY	P55246 oncorhynch
31	57.5	21.7	90	1 GON3_HAPBU	P45652 haplochromi
32	57.5	21.7	3133	1 HMCT_BOMMO	P98092 bombyx mori
33	57	21.5	309	1 FDHE_EC057	Q9x8B9 escherichia

ALIGNMENTS

RESULT 1

ID	CSP_PLARE	STANDARD;	PRT;	388 AA.
AC	P26694;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Circumsporozoite protein precursor (CS)			
OS	Plasmodium reichenowi.			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
OX	NCBI_TaxID=5854;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91201303; PubMed=2016283;			
RA	Lal A.A., Goldman I.F.;			
RT	"Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malaria parasite evolutionarily related to the human malaria parasite Plasmodium falciparum.";			
RL	J. Biol. Chem. 266:6886-6889(1991).			
CC	-!- FUNCTION: The circumsporozoite protein is the immunodominant surface antigen on the sporozoite (the infective stage of the malaria parasite that is transmitted from the mosquito to the vertebrate host).			
CC	-!- MISCELLANEOUS: The C-terminal region is probably used for anchoring the protein to the cell membrane. The repeat sequences would be the surface antigen of the organism.			
CC	-!- SIMILARITY: Contains 1 TSP type-1 domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M60972; AAA29561.1; -			
DR	PIR; A39756; A39756.			
DR	InterPro; IPR03067; Crcmsprzoite.			
DR	InterPro; IPR000864; TSP1.			
DR	Pfam; PF00090; tsp_1; 1.			
DR	PRINTS; PR01303; Crcmsprzoite.			
DR	SMART; SM00209; TSP1; 1.			
DR	PROSITE; PS00092; TSP1; 1.			
KW	Malaria; Sporozoite; Repeat; Signal.			
FT	SIGNAL	1	16	PROBABLE.
FT	CHAIN	1	388	CIRCUMSPOROZOITE PROTEIN.
FT	DOMAIN	120	263	37 X 4 AA TANDEM REPEATS OF N-[AV] - [ND] - P.
FT	REPEAT	120	123	1.
FT	REPEAT	124	127	2.
FT	REPEAT	128	131	3.
FT	REPEAT	132	135	4.
FT	REPEAT	136	139	5.
FT	REPEAT	140	143	6.
FT	REPEAT	144	147	7.

Q9f8es escherichia
P13024 escherichia
Q9hyg2 pseudomonas
O1410 bacillus st
P29466 homo sapien
O77780 bos taurus
P51922 porichnys
Q9dd49 o prognado
Q8cvu7 escherichia
Q9p327 schizosacch
P41484 mycobacteri
Q93j15 streptomyce

```

FT REPEAT 148 151 8.
FT REPEAT 152 155 9.
FT REPEAT 156 159 10.
FT REPEAT 160 163 11.
FT REPEAT 164 167 12.
FT REPEAT 168 171 13.
FT REPEAT 172 175 14.
FT REPEAT 176 179 15.
FT REPEAT 180 183 16.
FT REPEAT 184 187 17.
FT REPEAT 188 191 18.
FT REPEAT 192 195 20.
FT REPEAT 196 199 21.
FT REPEAT 200 203 22.
FT REPEAT 204 207 23.
FT REPEAT 208 211 24.
FT REPEAT 212 215 25.
FT REPEAT 216 219 26.
FT REPEAT 220 223 27.
FT REPEAT 224 227 28.
FT REPEAT 228 231 29.
FT REPEAT 232 235 30.
FT REPEAT 236 239 31.
FT REPEAT 240 243 32.
FT REPEAT 244 247 33.
FT REPEAT 248 251 34.
FT REPEAT 252 255 35.
FT REPEAT 256 259 36.
FT REPEAT 260 263 37.
FT DOMAIN 313 366 TSP TYPE-1.
SQ SEQUENCE 388 AA; 42245 MW; C031EFEB2B35604 CRC64;

Query Match 39.6%; Score 105; DB 1; Length 388;
Best Local Similarity 58.5%; Pred. No. 9.1e-06;
Matches 24; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

Qy 7 LRPGSSGSLD-----EKKIAKMEKASSVFNVNSSG 39
Db 337 IRPGSAGPKDQLDYENDLEKKICKMEKCSVFNVNSSIG 377

RESULT 2
CSP_PLAFO
ID CSP PLAFO STANDARD; PRT; 397 AA.
AC P19597; 025798;
DT 01-FEB-1991 (Rel. 17, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5843;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89345189; PubMed=2668895;
RA Campbell J.R.;
RT "DNA sequence of the gene encoding a Plasmodium falciparum malaria
candidate vaccine antigen.";
EL Nucleic Acids Res. 17:5854-5854 (1989).
RN [2]
RP REVISIONS.
RA Campbell J.R.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92155298; PubMed=1346766;
RA Davis J.R., Cortese J.F., Herrington D.A., Murphy J.R., Clyde D.F.,
RA Thomas A.W., Baqar S., Cochran M.A., Thanassi J., Levine M.M.,
RA Hackett C.S.;
RT "Plasmodium falciparum: in vitro characterization and human
infectivity of a cloned line.";
RL Exp. Parasitol. 74:159-168 (1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89345189; PubMed=2671723;
RA Caspers P., Gentz R., Matile H., Pink J.R., Sinisaglia P.;
RT "The circumsporozoite protein gene from NF54, a Plasmodium falciparum
isolate used in malaria vaccine trials.";
Mol. Biochem. Parasitol. 35:185-190 (1989).
RL
CC !- FUNCTION: The circumsporozoite protein is the immunodominant
surface antigen on the sporozoite (the infective stage of the
malaria parasite that is transmitted from the mosquito to the
vertebrate host).
CC
CC !- MISCELLANEOUS: The C-terminal region is probably used for
anchoring the protein to the cell membrane. The repeat sequences
would be the surface antigen of the organism.
CC
CC !- SIMILARITY: Contains 1 TSP type-1 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X15363; CAA33421.1; -
CC EMBL; M83866; AAA29521.1; -
CC EMBL; M22982; AAA29527.1; -
CC PIR; S05428; S05428.
CC InterPro; IPR003067; Crcmsprzoite.
CC Pfam; PF00090; tsp.1;
CC PRINTS; P801303; CRCMSPRZOITE.
CC SMART; SMO0209; TSP1; 1.
DR PROSITE; PS50052; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 397 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 105 272 42 X 4 AA TANDEM REPEATS OF N-[AV]-[ND]-
P.
FT REPEAT 105 108 1.
FT REPEAT 109 112 2.
FT REPEAT 113 116 3.
FT REPEAT 117 120 4.
FT REPEAT 121 124 5.
FT REPEAT 125 128 6.
FT REPEAT 129 132 7.
FT REPEAT 133 136 8.
FT REPEAT 137 140 9.
FT REPEAT 141 144 10.
FT REPEAT 145 148 11.
FT REPEAT 149 152 12.
FT REPEAT 153 156 13.
FT REPEAT 157 160 14.
FT REPEAT 161 164 15.
FT REPEAT 165 168 16.
FT REPEAT 169 172 17.
FT REPEAT 173 176 18.
FT REPEAT 177 180 19.
FT REPEAT 181 184 20.
FT REPEAT 185 188 21.
FT REPEAT 189 192 22.
FT REPEAT 193 196 23.
FT REPEAT 197 200 24.
FT REPEAT 201 204 25.
FT REPEAT 205 208 26.
FT REPEAT 209 212 27.
FT REPEAT 213 216 28.
FT REPEAT 217 220 29.
FT REPEAT 221 224 30.
FT REPEAT 225 228 31.
FT REPEAT 229 232 32.
FT REPEAT 233 236 33.
FT REPEAT 237 240 34.
FT REPEAT 241 244 35.

```

```

FT REPEAT 245 248 36.
FT REPEAT 249 252 37.
FT REPEAT 253 256 38.
FT REPEAT 257 260 39.
FT REPEAT 261 264 40.
FT REPEAT 265 268 41.
FT REPEAT 269 272 42.
FT DOMAIN 322 375 TSP TYPE-1.
FT CONFLICT 194 194 A -> ANPNANPNA (IN REF. 4).
SQ SEQUENCE 337 AA; 42646 MW; 9E81146F59EBCEA3 CRC64;

Query Match 37.4%; Score 99; DB 1; Length 397;
Best Local Similarity 56.1%; Pred. No. 5.5e-05;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAKWEKASSVFNVNSSG 39
DB 346 IKPGSANKPKDELVDYNDIEKKICKMEKCSVFNVNSSIG 386

RESULT 3
CSP PLAF A
ID - CSP PLAF A STANDARD; PRT; 412 AA.
AC P02893;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A. PubMed=6204383;
RA MEDLINE=84250215; PubMed=6204383;
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Hochmeyer W.T., Maloy J.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy E.P., Digs C.L., Miller L.H.;
RA "Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";
RL Science 225:593-599(1994).
CC -!- FUNCTION: The circumsporozoite protein is the immunodominant
CC surface antigen on the sporozoite (the infective stage of the
CC malaria parasite that is transmitted from the mosquito to the
CC vertebrate host).
CC -!- MISCELLANEOUS: The C-terminal region is probably used for
CC anchoring the protein to the cell membrane. The repeat sequences
CC would be the surface antigen of the organism.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02194; AAA29524.1; -.
DR PIR; A03388; OZZOAF.
DR InterPro; IPR003067; Crmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tep_1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS00092; TSP1; 1.
KW Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 16 PROBABLE.
FT CHAIN 17 412 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 123 286 41 X 4 AA TANDEM REPEATS OF P-N-[AV]-
FT REPEAT 123 126 [ND].
FT REPEAT 127 130 1.
FT REPEAT 131 134 2.
FT REPEAT 131 134 3.

```

```

FT REPEAT 135 138 4.
FT REPEAT 139 142 5.
FT REPEAT 143 146 6.
FT REPEAT 147 150 7.
FT REPEAT 151 154 8.
FT REPEAT 155 158 9.
FT REPEAT 159 162 10.
FT REPEAT 163 166 11.
FT REPEAT 167 170 12.
FT REPEAT 171 174 13.
FT REPEAT 175 178 14.
FT REPEAT 179 182 15.
FT REPEAT 183 186 16.
FT REPEAT 187 190 17.
FT REPEAT 191 194 18.
FT REPEAT 195 198 19.
FT REPEAT 199 202 20.
FT REPEAT 203 206 21.
FT REPEAT 207 210 22.
FT REPEAT 211 214 23.
FT REPEAT 215 218 24.
FT REPEAT 219 222 25.
FT REPEAT 223 226 26.
FT REPEAT 227 230 27.
FT REPEAT 231 234 28.
FT REPEAT 235 238 29.
FT REPEAT 239 242 30.
FT REPEAT 243 246 31.
FT REPEAT 247 250 32.
FT REPEAT 251 254 33.
FT REPEAT 255 258 34.
FT REPEAT 259 262 35.
FT REPEAT 263 266 36.
FT REPEAT 267 270 37.
FT REPEAT 271 274 38.
FT REPEAT 275 278 39.
FT REPEAT 279 282 40.
FT REPEAT 283 286 41.
FT DOMAIN 337 390 TSP TYPE-1.
SQ SEQUENCE 412 AA; 44420 MW; 1EEED3DB90965F8 CRC64;

Query Match 37.4%; Score 99; DB 1; Length 412;
Best Local Similarity 56.1%; Pred. No. 5.7e-05;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAKWEKASSVFNVNSSG 39
DB 361 IKPGSANKPKDELVDYNDIEKKICKMEKCSVFNVNSSIG 401

RESULT 4
CSP PLAF A
ID - CSP PLAF A STANDARD; PRT; 424 AA.
AC P13814;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium falciparum (isolate t4 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5846;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87315205; PubMed=3306373;
RA del Portillo H.A., Nussenzweig R.S., Enea V.;
RA "Circumsporozoite gene of a Plasmodium falciparum strain from
RT Thailand.";
EL Mol. Biochem. Parasitol. 24:289-294(1987).
CC -!- FUNCTION: The circumsporozoite protein is the immunodominant
CC surface antigen on the sporozoite (the infective stage of the
CC malaria parasite that is transmitted from the mosquito to the
CC vertebrate host).
CC -!- MISCELLANEOUS: The C-terminal region is probably used for

```

anchoring the protein to the cell membrane. The repeat sequences would be the surface antigen of the organism.

-!- SIMILARITY: Contains 1 TSP type-1 domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M19752; AAA29555.1; --

PIR; A54533; A54533.

InterPro; IPR003067; Crcmsprzoite.

InterPro; IPR000884; TSP1.

Pfam; PF00090; tsp 1; 1.

PRINTS; PR01303; CRCMSRZOITE.

SMART; SM00209; TSP1; 1.

PROSITE; PS50092; TSP1; 1.

Malaria; Sporozoite; Repeat; Signal.

PROBLE.

SIGNAL 1 16

CIRCUMSPOROZOITE PROTEIN.

44 X 4 AA TANDEM REPEATS OF P-N-[AV] -

[ND].

1. 126

2. 127

3. 130

4. 134

5. 138

6. 142

7. 146

8. 150

9. 154

10. 158

11. 162

12. 166

13. 170

14. 174

15. 178

16. 182

17. 186

18. 190

19. 194

20. 198

21. 202

22. 206

23. 210

24. 214

25. 218

26. 222

27. 226

28. 230

29. 234

30. 238

31. 242

32. 246

33. 250

34. 254

35. 258

36. 262

37. 266

38. 270

39. 274

40. 278

41. 282

42. 286

43. 290

44. 294

45. 298

46. 302

47. 306

48. 310

49. 314

50. 318

51. 322

52. 326

53. 330

54. 334

55. 338

56. 342

57. 346

58. 350

59. 354

60. 358

61. 362

62. 366

63. 370

64. 374

65. 378

66. 382

67. 386

68. 390

69. 394

70. 398

71. 402

72. 406

73. 410

74. 414

75. 418

76. 422

77. 426

78. 430

79. 434

80. 438

81. 442

82. 446

83. 450

84. 454

85. 458

86. 462

87. 466

88. 470

89. 474

90. 478

91. 482

92. 486

93. 490

94. 494

95. 498

96. 502

97. 506

98. 510

99. 514

100. 518

101. 522

102. 526

103. 530

104. 534

105. 538

106. 542

107. 546

108. 550

109. 554

110. 558

111. 562

112. 566

113. 570

114. 574

115. 578

116. 582

117. 586

118. 590

119. 594

120. 598

121. 602

122. 606

123. 610

124. 614

125. 618

126. 622

127. 626

128. 630

129. 634

130. 638

131. 642

132. 646

133. 650

134. 654

135. 658

136. 662

137. 666

138. 670

139. 674

140. 678

141. 682

142. 686

143. 690

144. 694

145. 698

146. 702

147. 706

148. 710

149. 714

150. 718

151. 722

152. 726

153. 730

154. 734

155. 738

156. 742

157. 746

158. 750

159. 754

160. 758

161. 762

162. 766

163. 770

164. 774

165. 778

166. 782

167. 786

168. 790

169. 794

170. 798

171. 802

172. 806

173. 810

174. 814

175. 818

176. 822

177. 826

178. 830

179. 834

180. 838

181. 842

182. 846

183. 850

184. 854

185. 858

186. 862

187. 866

188. 870

189. 874

190. 878

191. 882

192. 886

193. 890

194. 894

195. 898

196. 902

197. 906

198. 910

199. 914

200. 918

201. 922

202. 926

203. 930

204. 934

205. 938

206. 942

207. 946

208. 950

209. 954

210. 958

211. 962

212. 966

213. 970

214. 974

215. 978

216. 982

217. 986

218. 990

219. 994

220. 998

221. 1002

222. 1006

223. 1010

224. 1014

225. 1018

226. 1022

227. 1026

228. 1030

229. 1034

230. 1038

231. 1042

232. 1046

233. 1050

234. 1054

235. 1058

236. 1062

237. 1066

238. 1070

239. 1074

240. 1078

241. 1082

242. 1086

243. 1090

244. 1094

245. 1098

246. 1102

247. 1106

248. 1110


```

FT REPEAT 202 205 18.
FT REPEAT 206 209 19.
FT REPEAT 210 213 20.
FT REPEAT 214 217 21.
FT REPEAT 218 221 22.
FT REPEAT 222 225 23.
FT REPEAT 226 229 24.
FT REPEAT 230 233 25.
FT REPEAT 234 237 26.
FT REPEAT 238 241 27.
FT REPEAT 242 245 28.
FT REPEAT 246 249 29.
FT REPEAT 250 253 30.
FT REPEAT 254 257 31.
FT REPEAT 258 261 32.
FT REPEAT 262 265 33.
FT REPEAT 266 269 34.
FT REPEAT 270 273 35.
FT REPEAT 274 277 36.
FT REPEAT 282 285 38.
FT REPEAT 286 289 39.
FT REPEAT 290 293 40.
FT REPEAT 294 297 41.
FT REPEAT 298 301 42.
FT REPEAT 302 305 43.
FT REPEAT 306 309 44.
FT REPEAT 310 313 45.
FT REPEAT 314 317 46.
FT DOMAIN 367 420 TSP TYPE-1.
SQ SEQUENCE 442 AA; 47402 MW; BD57A9A152B85E03 CRC64;

Query Match 37.0%; Score 98; DB 1; Length 442;
Best Local Similarity 56.1%; Pred. No. 8.3e-05;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGSLD-----EKKIAKMEKASVFNVNSSG 39
DQ 391 IKPGSADPKQDLYENDIERKICKMEKCSSVFNVNSSIG 431

RESULT 6
CSP_PLABE STANDARD; PRT; 339 AA.
AC P06315;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium berghei.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5821;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87089740; PubMed=2432395;
RA Eichinger D.J., Arnot D.E., Tam J.P., Nussenzweig V., Enea V.;
RT "Circumsporozoite protein of Plasmodium berghei: gene cloning and
RT identification of the immunodominant epitopes.";
RL Mol. Cell. Biol. 6:3965-3972(1986).
CC -!- FUNCTION: The circumsporozoite protein is the immunodominant
CC surface antigen on the sporozoite (the infective stage of the
CC malaria parasite that is transmitted from the mosquito to the
CC vertebrate host).
CC -!- MISCELLANEOUS: The C-terminal region is probably used for
CC anchoring the protein to the cell membrane. The repeat sequences
CC would be the surface antigen of the organism.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M14135; AAA29577.1; -.
CC PIR; A44948; OZZOMB.
CC InterPro; IPR003067; Crcmsprzoite.
CC InterPro; IPR000884; TSP1.
CC Pfam; PF00900; tsp_1; 1.
CC PRINTS; PRO1303; Crcmsprzoite.
CC SMART; SM00209; TSP1; 1.
CC PROSITE; PS00092; TSP1; 1.
CC Malaria; Sporozoite; Repeat; Signal.
FT SIGNAL 1 23 PROBABLE.
FT CHAIN 24 339 CIRCUMSPOROZOITE PROTEIN.
FT DOMAIN 53 136 13 X 8 AA TANDEM REPEATS OF P-[PA]-P-P-N-
FT REPEAT 93 100 [PA]-N-D.
FT REPEAT 101 108 1-1.
FT REPEAT 109 116 1-2.
FT REPEAT 117 124 1-3.
FT REPEAT 125 132 1-4.
FT REPEAT 133 140 1-5.
FT REPEAT 141 148 1-6.
FT REPEAT 149 156 1-7.
FT REPEAT 157 164 1-8.
FT REPEAT 165 172 1-9.
FT REPEAT 173 180 1-10.
FT REPEAT 181 188 1-11.
FT REPEAT 189 196 1-12.
FT DOMAIN 206 238 1-13.
FT REPEAT 206 207 17 X 2 AA APPROXIMATE TANDEM REPEATS OF
FT REPEAT 208 209 P-Q.
FT REPEAT 210 211 2-1.
FT REPEAT 212 213 2-2.
FT REPEAT 214 215 2-3 (APPROXIMATE).
FT REPEAT 216 217 2-4.
FT REPEAT 218 219 2-5.
FT REPEAT 220 221 2-6.
FT REPEAT 222 223 2-7.
FT REPEAT 224 225 2-8.
FT REPEAT 226 227 2-9.
FT REPEAT 228 229 2-10.
FT REPEAT 230 231 2-11.
FT REPEAT 232 233 2-12.
FT REPEAT 234 235 2-13 (APPROXIMATE).
FT REPEAT 236 237 2-14.
FT REPEAT 238 239 2-15.
FT DOMAIN 266 317 TSP TYPE-1.
SQ SEQUENCE 339 AA; 37138 MW; E8068A6D11D9551B CRC64;

Query Match 24.9%; Score 66; DB 1; Length 339;
Best Local Similarity 39.5%; Pred. No. 0.73;
Matches 15; Conservative 9; Mismatches 8; Indels 6; Gaps 1;

QY 8 RGSSGSLD-----EKKIAKMEKASVFNVNSSG 39
DQ 291 RKGSKKAEDLTEDITEICKDKCSSIFNIVNSLG 328

RESULT 7
CSP_PLABE STANDARD; PRT; 347 AA.
AC P23033;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Circumsporozoite protein precursor (CS).
OS Plasmodium berghei (strain Anka).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5823;
RN [1]
RP SEQUENCE FROM N.A.

```

RX MEDLINE=90221834; PubMed=2183186;
 RA Lockyer M.J., Davies C.S., Shribier A., Sinden R.E.;
 RT "Nucleotide sequence of the Plasmodium berghei circumsporozoite
 RL Nucleic Acids Res. 18:376-376(1990).
 CC -!- FUNCTION: The circumsporozoite protein is the immunodominant
 CC surface antigen on the sporozoite (the infective stage of the
 CC malaria parasite that is transmitted from the mosquito to the
 CC vertebrate host).
 CC -!- MISCELLANEOUS: The C-terminal region is probably used for
 CC anchoring the protein to the cell membrane. The repeat sequences
 CC would be the surface antigen of the organism.
 CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X17606; CAA35608.1; -;
 DR PIR; S07873; OZZOBK.
 DR InterPro; IPR003067; Crmsprzoite.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PR01303; CRMSPRZOITE.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PS00092; TSP1; 1.
 KW Malaria; Sporozoite; Repeat; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 347
 FT DOMAIN 93 196
 FT REPEAT 93 100
 FT REPEAT 101 108
 FT REPEAT 109 116
 FT REPEAT 117 124
 FT REPEAT 125 132
 FT REPEAT 133 140
 FT REPEAT 141 148
 FT REPEAT 149 156
 FT REPEAT 157 164
 FT REPEAT 165 172
 FT REPEAT 173 180
 FT REPEAT 181 188
 FT REPEAT 189 196
 FT DOMAIN 214 247
 FT REPEAT 214 215
 FT REPEAT 216 217
 FT REPEAT 218 219
 FT REPEAT 220 221
 FT REPEAT 222 223
 FT REPEAT 224 225
 FT REPEAT 226 227
 FT REPEAT 228 229
 FT REPEAT 230 231
 FT REPEAT 232 233
 FT REPEAT 234 235
 FT REPEAT 236 237
 FT REPEAT 238 239
 FT REPEAT 240 241
 FT REPEAT 242 243
 FT REPEAT 244 245
 FT REPEAT 246 247
 FT DOMAIN 274 325
 SQ SEQUENCE 347 AA; 37776 MW; 0EC240E35681AF8 CRC64;
 Query Match 24.9%; Score 66; DB 1; Length 347;
 Best Local Similarity 39.5%; Pred. No. 0.75;
 Matches 15; Conservative 9; Mismatches 8; Indels 6; Gaps 1;

QY 8 RFGSGPSLD-----EKKIAMEKASSVFVNVSNSG 39
 DB 299 RKGSKKAEDLTLEDIDTEICKMKCKSSIFIVNSLGL 336
 RESULT 8
 CSP PLAYO STANDARD; PRT; 367 AA.
 ID CSP PLAYO
 AC P06314;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Circumsporozoite protein precursor (CS).
 OS Plasmodium berghei yoelii.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5862;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87137555; PubMed=3102479;
 RA Lal A.A., de la Cruz V.F., Welsh J.A., Charoenvit Y., Maloy W.L.,
 RA McCutchan T.F.;
 RT "Structure of the gene encoding the circumsporozoite protein of
 RT Plasmodium yoelii. A rodent model for examining antimalarial
 RT sporozoite vaccines";
 RL J. Biol. Chem. 262:2937-2940(1987).
 RN [2]
 RP SEQUENCE OF 1-140 AND 260-367 FROM N.A.
 RX MEDLINE=88232798; PubMed=3287156;
 RA de la Cruz V.F., Lal A.A., McCutchan T.F.;
 RT "Variation among circumsporozoite protein genes from rodent
 RT malarial parasites";
 RL Mol. Biochem. Parasitol. 28:31-38(1988).
 CC -!- FUNCTION: The circumsporozoite protein is the immunodominant
 CC surface antigen on the sporozoite (the infective stage of the
 CC malaria parasite that is transmitted from the mosquito to the
 CC vertebrate host).
 CC -!- MISCELLANEOUS: The C-terminal region is probably used for
 CC anchoring the protein to the cell membrane. The repeat sequences
 CC would be the surface antigen of the organism.
 CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; J02695; AAA29558.1; -;
 DR EMBL; M18821; AAA29559.1; -;
 DR EMBL; M22698; AAA29560.1; -;
 DR InterPro; IPR003067; Crmsprzoite.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; tsp_1; 1.
 DR PRINTS; PR01303; CRMSPRZOITE.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PS00092; TSP1; 1.
 KW Malaria; Sporozoite; Repeat; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 367
 FT DOMAIN 139 228
 FT REPEAT 139 144
 FT REPEAT 145 150
 FT REPEAT 151 156
 FT REPEAT 157 162
 FT REPEAT 163 168
 FT REPEAT 169 174
 FT REPEAT 175 180
 FT REPEAT 181 186
 FT REPEAT 187 192
 FT REPEAT 193 198
 FT REPEAT 199 200
 FT REPEAT 201 202
 FT REPEAT 203 204
 FT REPEAT 205 206
 FT REPEAT 207 208
 FT REPEAT 209 210
 FT REPEAT 211 212
 FT REPEAT 213 214
 FT REPEAT 215 216
 FT REPEAT 217 218
 FT REPEAT 219 220
 FT REPEAT 221 222
 FT REPEAT 223 224
 FT REPEAT 225 226
 FT REPEAT 227 228
 FT REPEAT 229 230
 FT REPEAT 231 232
 FT REPEAT 233 234
 FT REPEAT 235 236
 FT REPEAT 237 238
 FT REPEAT 239 240
 FT REPEAT 241 242
 FT REPEAT 243 244
 FT REPEAT 245 246
 FT REPEAT 247 248
 FT REPEAT 249 250
 FT REPEAT 251 252
 FT REPEAT 253 254
 FT REPEAT 255 256
 FT REPEAT 257 258
 FT REPEAT 259 260
 FT REPEAT 261 262
 FT REPEAT 263 264
 FT REPEAT 265 266
 FT REPEAT 267 268
 FT REPEAT 269 270
 FT REPEAT 271 272
 FT REPEAT 273 274
 FT REPEAT 275 276
 FT REPEAT 277 278
 FT REPEAT 279 280
 FT REPEAT 281 282
 FT REPEAT 283 284
 FT REPEAT 285 286
 FT REPEAT 287 288
 FT REPEAT 289 290
 FT REPEAT 291 292
 FT REPEAT 293 294
 FT REPEAT 295 296
 FT REPEAT 297 298
 FT REPEAT 299 300
 FT REPEAT 301 302
 FT REPEAT 303 304
 FT REPEAT 305 306
 FT REPEAT 307 308
 FT REPEAT 309 310
 FT REPEAT 311 312
 FT REPEAT 313 314
 FT REPEAT 315 316
 FT REPEAT 317 318
 FT REPEAT 319 320
 FT REPEAT 321 322
 FT REPEAT 323 324
 FT REPEAT 325 326
 FT REPEAT 327 328
 FT REPEAT 329 330
 FT REPEAT 331 332
 FT REPEAT 333 334
 FT REPEAT 335 336
 FT REPEAT 337 338
 FT REPEAT 339 340
 FT REPEAT 341 342
 FT REPEAT 343 344
 FT REPEAT 345 346
 FT REPEAT 347 348
 FT REPEAT 349 350
 FT REPEAT 351 352
 FT REPEAT 353 354
 FT REPEAT 355 356
 FT REPEAT 357 358
 FT REPEAT 359 360
 FT REPEAT 361 362
 FT REPEAT 363 364
 FT REPEAT 365 366
 FT REPEAT 367 368
 FT REPEAT 369 370
 FT REPEAT 371 372
 FT REPEAT 373 374
 FT REPEAT 375 376
 FT REPEAT 377 378
 FT REPEAT 379 380
 FT REPEAT 381 382
 FT REPEAT 383 384
 FT REPEAT 385 386
 FT REPEAT 387 388
 FT REPEAT 389 390
 FT REPEAT 391 392
 FT REPEAT 393 394
 FT REPEAT 395 396
 FT REPEAT 397 398
 FT REPEAT 399 400
 FT REPEAT 401 402
 FT REPEAT 403 404
 FT REPEAT 405 406
 FT REPEAT 407 408
 FT REPEAT 409 410
 FT REPEAT 411 412
 FT REPEAT 413 414
 FT REPEAT 415 416
 FT REPEAT 417 418
 FT REPEAT 419 420
 FT REPEAT 421 422
 FT REPEAT 423 424
 FT REPEAT 425 426
 FT REPEAT 427 428
 FT REPEAT 429 430
 FT REPEAT 431 432
 FT REPEAT 433 434
 FT REPEAT 435 436
 FT REPEAT 437 438
 FT REPEAT 439 440
 FT REPEAT 441 442
 FT REPEAT 443 444
 FT REPEAT 445 446
 FT REPEAT 447 448
 FT REPEAT 449 450
 FT REPEAT 451 452
 FT REPEAT 453 454
 FT REPEAT 455 456
 FT REPEAT 457 458
 FT REPEAT 459 460
 FT REPEAT 461 462
 FT REPEAT 463 464
 FT REPEAT 465 466
 FT REPEAT 467 468
 FT REPEAT 469 470
 FT REPEAT 471 472
 FT REPEAT 473 474
 FT REPEAT 475 476
 FT REPEAT 477 478
 FT REPEAT 479 480
 FT REPEAT 481 482
 FT REPEAT 483 484
 FT REPEAT 485 486
 FT REPEAT 487 488
 FT REPEAT 489 490
 FT REPEAT 491 492
 FT REPEAT 493 494
 FT REPEAT 495 496
 FT REPEAT 497 498
 FT REPEAT 499 500
 FT REPEAT 501 502
 FT REPEAT 503 504
 FT REPEAT 505 506
 FT REPEAT 507 508
 FT REPEAT 509 510
 FT REPEAT 511 512
 FT REPEAT 513 514
 FT REPEAT 515 516
 FT REPEAT 517 518
 FT REPEAT 519 520
 FT REPEAT 521 522
 FT REPEAT 523 524
 FT REPEAT 525 526
 FT REPEAT 527 528
 FT REPEAT 529 530
 FT REPEAT 531 532
 FT REPEAT 533 534
 FT REPEAT 535 536
 FT REPEAT 537 538
 FT REPEAT 539 540
 FT REPEAT 541 542
 FT REPEAT 543 544
 FT REPEAT 545 546
 FT REPEAT 547 548
 FT REPEAT 549 550
 FT REPEAT 551 552
 FT REPEAT 553 554
 FT REPEAT 555 556
 FT REPEAT 557 558
 FT REPEAT 559 560
 FT REPEAT 561 562
 FT REPEAT 563 564
 FT REPEAT 565 566
 FT REPEAT 567 568
 FT REPEAT 569 570
 FT REPEAT 571 572
 FT REPEAT 573 574
 FT REPEAT 575 576
 FT REPEAT 577 578
 FT REPEAT 579 580
 FT REPEAT 581 582
 FT REPEAT 583 584
 FT REPEAT 585 586
 FT REPEAT 587 588
 FT REPEAT 589 590
 FT REPEAT 591 592
 FT REPEAT 593 594
 FT REPEAT 595 596
 FT REPEAT 597 598
 FT REPEAT 599 600
 FT REPEAT 601 602
 FT REPEAT 603 604
 FT REPEAT 605 606
 FT REPEAT 607 608
 FT REPEAT 609 610
 FT REPEAT 611 612
 FT REPEAT 613 614
 FT REPEAT 615 616
 FT REPEAT 617 618
 FT REPEAT 619 620
 FT REPEAT 621 622
 FT REPEAT 623 624
 FT REPEAT 625 626
 FT REPEAT 627 628
 FT REPEAT 629 630
 FT REPEAT 631 632
 FT REPEAT 633 634
 FT REPEAT 635 636
 FT REPEAT 637 638
 FT REPEAT 639 640
 FT REPEAT 641 642
 FT REPEAT 643 644
 FT REPEAT 645 646
 FT REPEAT 647 648
 FT REPEAT 649 650
 FT REPEAT 651 652
 FT REPEAT 653 654
 FT REPEAT 655 656
 FT REPEAT 657 658
 FT REPEAT 659 660
 FT REPEAT 661 662
 FT REPEAT 663 664
 FT REPEAT 665 666
 FT REPEAT 667 668
 FT REPEAT 669 670
 FT REPEAT 671 672
 FT REPEAT 673 674
 FT REPEAT 675 676
 FT REPEAT 677 678
 FT REPEAT 679 680
 FT REPEAT 681 682
 FT REPEAT 683 684
 FT REPEAT 685 686
 FT REPEAT 687 688
 FT REPEAT 689 690
 FT REPEAT 691 692
 FT REPEAT 693 694
 FT REPEAT 695 696
 FT REPEAT 697 698
 FT REPEAT 699 700
 FT REPEAT 701 702
 FT REPEAT 703 704
 FT REPEAT 705 706
 FT REPEAT 707 708
 FT REPEAT 709 710
 FT REPEAT 711 712
 FT REPEAT 713 714
 FT REPEAT 715 716
 FT REPEAT 717 718
 FT REPEAT 719 720
 FT REPEAT 721 722
 FT REPEAT 723 724
 FT REPEAT 725 726
 FT REPEAT 727 728
 FT REPEAT 729 730
 FT REPEAT 731 732
 FT REPEAT 733 734
 FT REPEAT 735 736
 FT REPEAT 737 738
 FT REPEAT 739 740
 FT REPEAT 741 742
 FT REPEAT 743 744
 FT REPEAT 745 746
 FT REPEAT 747 748
 FT REPEAT 749 750
 FT REPEAT 751 752
 FT REPEAT 753 754
 FT REPEAT 755 756
 FT REPEAT 757 758
 FT REPEAT 759 760
 FT REPEAT 761 762
 FT REPEAT 763 764
 FT REPEAT 765 766
 FT REPEAT 767 768
 FT REPEAT 769 770
 FT REPEAT 771 772
 FT REPEAT 773 774
 FT REPEAT 775 776
 FT REPEAT 777 778
 FT REPEAT 779 780
 FT REPEAT 781 782
 FT REPEAT 783 784
 FT REPEAT 785 786
 FT REPEAT 787 788
 FT REPEAT 789 790
 FT REPEAT 791 792
 FT REPEAT 793 794
 FT REPEAT 795 796
 FT REPEAT 797 798
 FT REPEAT 799 800
 FT REPEAT 801 802
 FT REPEAT 803 804
 FT REPEAT 805 806
 FT REPEAT 807 808
 FT REPEAT 809 810
 FT REPEAT 811 812
 FT REPEAT 813 814
 FT REPEAT 815 816
 FT REPEAT 817 818
 FT REPEAT 819 820
 FT REPEAT 821 822
 FT REPEAT 823 824
 FT REPEAT 825 826
 FT REPEAT 827 828
 FT REPEAT 829 830
 FT REPEAT 831 832
 FT REPEAT 833 834
 FT REPEAT 835 836
 FT REPEAT 837 838
 FT REPEAT 839 840
 FT REPEAT 841 842
 FT REPEAT 843 844
 FT REPEAT 845 846
 FT REPEAT 847 848
 FT REPEAT 849 850
 FT REPEAT 851 852
 FT REPEAT 853 854
 FT REPEAT 855 856
 FT REPEAT 857 858
 FT REPEAT 859 860
 FT REPEAT 861 862
 FT REPEAT 863 864
 FT REPEAT 865 866
 FT REPEAT 867 868
 FT REPEAT 869 870
 FT REPEAT 871 872
 FT REPEAT 873 874
 FT REPEAT 875 876
 FT REPEAT 877 878
 FT REPEAT 879 880
 FT REPEAT 881 882
 FT REPEAT 883 884
 FT REPEAT 885 886
 FT REPEAT 887 888
 FT REPEAT 889 890
 FT REPEAT 891 892
 FT REPEAT 893 894
 FT REPEAT 895 896
 FT REPEAT 897 898
 FT REPEAT 899 900
 FT REPEAT 901 902
 FT REPEAT 903 904
 FT REPEAT 905 906
 FT REPEAT 907 908
 FT REPEAT 909 910
 FT REPEAT 911 912
 FT REPEAT 913 914
 FT REPEAT 915 916
 FT REPEAT 917 918
 FT REPEAT 919 920
 FT REPEAT 921 922
 FT REPEAT 923 924
 FT REPEAT 925 926
 FT REPEAT 927 928
 FT REPEAT 929 930
 FT REPEAT 931 932
 FT REPEAT 933 934
 FT REPEAT 935 936
 FT REPEAT 937 938
 FT REPEAT 939 940
 FT REPEAT 941 942
 FT REPEAT 943 944
 FT REPEAT 945 946
 FT REPEAT 947 948
 FT REPEAT 949 950
 FT REPEAT 951 952
 FT REPEAT 953 954
 FT REPEAT 955 956
 FT REPEAT 957 958
 FT REPEAT 959 960
 FT REPEAT 961 962
 FT REPEAT 963 964
 FT REPEAT 965 966
 FT REPEAT 967 968
 FT REPEAT 969 970
 FT REPEAT 971 972
 FT REPEAT 973 974
 FT REPEAT 975 976
 FT REPEAT 977 978
 FT REPEAT 979 980
 FT REPEAT 981 982
 FT REPEAT 983 984
 FT REPEAT 985 986
 FT REPEAT 987 988
 FT REPEAT 989 990
 FT REPEAT 991 992
 FT REPEAT 993 994
 FT REPEAT 995 996
 FT REPEAT 997 998
 FT REPEAT 999 1000
 FT REPEAT 1001 1002
 FT REPEAT 1003 1004
 FT REPEAT 1005 1006
 FT REPEAT 1007 1008
 FT REPEAT 1009 1010
 FT REPEAT 1011 1012
 FT REPEAT 1013 1014
 FT REPEAT 1015 1016
 FT REPEAT 1017 1018
 FT REPEAT 1019 1020
 FT REPEAT 1021 1022
 FT REPEAT 1023 1024
 FT REPEAT 1025 1026
 FT REPEAT 1027 1028
 FT REPEAT 1029 1030
 FT REPEAT 1031 1032
 FT REPEAT 1033 1034
 FT REPEAT 1035 1036
 FT REPEAT 1037 1038
 FT REPEAT 1039 1040
 FT REPEAT 1041 1042
 FT REPEAT 1043 1044
 FT REPEAT 1045 1046
 FT REPEAT 1047 1048
 FT REPEAT 1049 1050
 FT REPEAT 1051 1052
 FT REPEAT 1053 1054
 FT REPEAT 1055 1056
 FT REPEAT 1057 1058
 FT REPEAT 1059 1060
 FT REPEAT 1061 1062
 FT REPEAT 1063 1064
 FT REPEAT 1065 1066
 FT REPEAT 1067 1068
 FT REPEAT 1069 1070
 FT REPEAT 1071 1072
 FT REPEAT 1073 1074
 FT REPEAT 1075 1076
 FT REPEAT 1077 1078
 FT REPEAT 1079 1080
 FT REPEAT 1081 1082
 FT REPEAT 1083 1084
 FT REPEAT 1085 1086
 FT REPEAT 1087 1088
 FT REPEAT 1089 1090
 FT REPEAT 1091 1092
 FT REPEAT 1093 1094
 FT REPEAT 1095 1096
 FT REPEAT 1097 1098
 FT REPEAT 1099 1100
 FT REPEAT 1101 1102
 FT REPEAT 1103 1104
 FT REPEAT 1105 1106
 FT REPEAT 1107 1108
 FT REPEAT 1109 1110
 FT REPEAT 1111 1112
 FT REPEAT 1113 1114
 FT REPEAT 1115 1116
 FT REPEAT 1117 1118
 FT REPEAT 1119 1120
 FT REPEAT 1121 1122
 FT REPEAT 1123 1124
 FT REPEAT 1125 1126
 FT REPEAT 1127 1128
 FT REPEAT 1129 1130
 FT REPEAT 1131 1132
 FT REPEAT 1133 1134
 FT REPEAT 1135 1136
 FT REPEAT 1137 1138
 FT REPEAT 1139 1140
 FT REPEAT 1141 1142
 FT REPEAT 1143 1144
 FT REPEAT 1145 1146
 FT REPEAT 1147 1148
 FT REPEAT 1149 1150
 FT REPEAT 1151 1152
 FT REPEAT 1153 1154
 FT REPEAT 1155 1156
 FT REPEAT 1157 1158
 FT REPEAT 1159 1160
 FT REPEAT 1161 1162
 FT REPEAT 1163 1164
 FT REPEAT 1165 1166
 FT REPEAT 1167 1168
 FT REPEAT 1169 1170
 FT REPEAT 1171 1172
 FT REPEAT 1173 1174
 FT REPEAT 1175 1176
 FT REPEAT 1177 1178
 FT REPEAT 1179 1180
 FT REPEAT 1181 1182
 FT REPEAT 1183 1184
 FT REPEAT 1185 1186
 FT REPEAT 1187 1188
 FT REPEAT 1189 1190
 FT REPEAT 1191 1192
 FT REPEAT 1193 1194
 FT REPEAT 1195 1196
 FT REPEAT 1197 1198
 FT REPEAT 1199 1200
 FT REPEAT 1201 1202
 FT REPEAT 1203 1204
 FT REPEAT 1205 1206
 FT REPEAT 1207 1208
 FT REPEAT 1209 1210
 FT REPEAT 1211 1212
 FT REPEAT 1213 1214
 FT REPEAT 1215 1216
 FT REPEAT 1217 1218
 FT REPEAT 1219 1220
 FT REPEAT 1221 1222
 FT REPEAT 1223 1224
 FT REPEAT 1225 1226
 FT REPEAT 1227 1228
 FT REPEAT 1229 1230
 FT REPEAT 1231 1232
 FT REPEAT 1233 1234
 FT REPEAT 1235 1236
 FT REPEAT 1237 1238
 FT REPEAT 1239 1240
 FT REPEAT 1241 1242
 FT REPEAT 1243 1244
 FT REPEAT 1245 1246
 FT REPEAT 1247 1248
 FT REPEAT 1249 1250
 FT REPEAT 1251 1252
 FT REPEAT 1253 1254
 FT REPEAT 1255 1256
 FT REPEAT 1257 1258
 FT REPEAT 1259 1260
 FT REPEAT 1261 1262
 FT REPEAT 1263 1264
 FT REPEAT 1265 1266
 FT REPEAT 1267 1268
 FT REPEAT 1269 1270
 FT REPEAT 1271 1272
 FT REPEAT 1273 1274
 FT REPEAT 1275 1276
 FT REPEAT 1277 1278
 FT REPEAT 1279 1280
 FT REPEAT 1281 1282
 FT REPEAT 1283 1284
 FT REPEAT 1285 1286
 FT REPEAT 1287 1288
 FT REPEAT 1289 1290
 FT REPEAT 1291 1292
 FT REPEAT 1293 1294
 FT REPEAT 1295 1296
 FT REPEAT 1297 1298
 FT REPEAT 1299 1300
 FT REPEAT 1301 1302
 FT REPEAT 1303 1304
 FT REPEAT 1305 1306
 FT REPEAT 1307 1308
 FT REPEAT 1309 1310
 FT REPEAT 1311 1312
 FT REPEAT 1313 1314
 FT REPEAT 1315 1316
 FT REPEAT 1317 1318
 FT REPEAT 1319 1320
 FT REPEAT 1321 1322
 FT REPEAT 1323 1324
 FT REPEAT 1325 1326
 FT REPEAT 1327 1328
 FT REPEAT 1329 1330
 FT REPEAT 1331 1332
 FT REPEAT 1333 1334
 FT REPEAT 1335 1336
 FT REPEAT 1337 1338
 FT REPEAT 1339 1340
 FT REPEAT 1341 1342
 FT REPEAT 1343 1344
 FT REPEAT 1345 1346
 FT REPEAT 1347 1348
 FT REPEAT 1349 1350
 FT REPEAT 1351 1352
 FT REPEAT 1353 1354
 FT REPEAT 1355 1356
 FT REPEAT 1357 1358
 FT REPEAT 1359 1360
 FT REPEAT 1361 1362
 FT REPEAT 1363 1364
 FT REPEAT 1365 1366
 FT REPEAT 1367 1368
 FT REPEAT 1369 1370
 FT REPEAT 1371 1372
 FT REPEAT 1373 1374
 FT REPEAT 1375 1376
 FT REPEAT 1377 1378
 FT REPEAT 1379 1380
 FT REPEAT 1381 1382
 FT REPEAT 1383 1384
 FT REPEAT 1385 1386
 FT REPEAT 1387 1388
 FT REPEAT 1389 1390
 FT REPEAT 1391 1392
 FT REPEAT 1393 1394
 FT REPEAT 1395 1396
 FT REPEAT 1397 1398
 FT REPEAT 1399 1400
 FT REPEAT 1401 1402
 FT REPEAT 1403 1404
 FT REPEAT 1405 1406
 FT REPEAT 1407 1408
 FT REPEAT 1409 1410
 FT REPEAT 1411 1412
 FT REPEAT 1413 1414
 FT REPEAT 1415 1416
 FT REPEAT 1417 1418
 FT REPEAT 1419 1420
 FT REPEAT 1421 1422
 FT REPEAT 1423 1424
 FT REPEAT 1425 1426
 FT REPEAT 1427 14

```

FT PEPTIDE 37 82 GNRH-ASSOCIATED PEPTIDE III (POTENTIAL) .
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) .
SQ SEQUENCE 82 AA; 9191 MW; 8053B9534A765408 CRC64;

Query Match 23.6%; Score 62.5; DB 1; Length 82;
Best Local Similarity 32.7%; Pred. No. 0.39; 20; Indels 9; Gaps 2;
Matches 17; Conservative 6; Mismatches 20;

QY 2 HWSYGLRPGSSPSLDEKIAKVE---KASSVFNVNSSGPSLHWSYGLRP 50
||| ||| : : : ||| : |||
: : : : : : : : :
Db 25 HWSYGLPGG-----KRSVGELEATIKWDTGGVVALPEETSAAHFSERLRP 70
||| ||| : : : ||| : |||
: : : : : : : : :

RESULT 10
GONI_MESAU
ID GONI_MESAU STANDARD; PRT; 63 AA.
AC 009163;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE (Fragment).
DE GNRH1 OR GNRH OR LHRH.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
OX [1]
RP SEQUENCE FROM N.A.
RA Jansen H.T., Stevens P.J., Zeitler P., Lehman M.N.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U91938; AAB51302.1; -.
CC InterPro; IPR002012; GnRH.
CC DR InterPro; IPR004079; Gonadoliberin1.
CC DR Pfam; PF00446; GnRH; 1.
CC DR PRINTS; PR01541; GONADOLIBRNI.
CC DR PROSITE; PS00473; GNRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Pyrrolidone carboxylic acid.
FT NON_TER 1 1
FT CHAIN 1 1
FT PEPTIDE 1 10
FT PEPTIDE 14 63
FT ACT_SITE 3 3
FT FT 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT FT 1 1 ACTIVITY (BY SIMILARITY).
FT FT 1 1 PYRROLIDONE CARBOXYLIC ACID (BY
FT FT 1 1 SIMILARITY).
FT FT 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP) (BY
FT FT 63 63 SIMILARITY).
FT NON_TER 63 63
FT SEQUENCE 63 AA; 7370 MW; FC94995676F77180 CRC64;

Query Match 23.4%; Score 62; DB 1; Length 63;
Best Local Similarity 33.3%; Pred. No. 0.33;

```

Matches 18; Conservative 5; Mismatches 9; Indels 22; Gaps 4;
 QY 2 HWSYGLRPGSS-----GSLDEKKIAKMEKASSVFVNSSSGP-----SLHW 44
 |||||
 Db 2 HWSYGLRPGGKRAERLGDSEFQ-----MDKE-----VDQLAEPQHLECTVHW 44
 |||||

RESULT 11
 GON3 SHEEP
 ID GON3 SHEEP STANDARD; PRT; 61 AA.
 AC Q28588;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadoliberein I precursor (Contains: Gonadoliberein I (LH-RH I)
 (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I)
 (Fragment).
 DE GNRH1 OR GNRH OR LHRH.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE OF 12-61 FROM N.A.
 RC STRAIN=Western range; TISSUE=Hypothalamus;
 RA Rodriguez R.E., Wise M.E.;
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-10.
 RX MEDLINE=72094314; PubMed=4550508;
 RA Burgess R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,
 Fellows R., Blackwell R., Vale W., Guillemin R.;
 RT "Primary structure of the ovine hypothalamic luteinizing hormone-
 releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass
 spectrometry-decapeptide-sdman degradation).";
 RL Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 the secretion of both luteinizing and follicle-stimulating
 hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnRH family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; U02517; AAA03433.1; -.
 DR InterPro; IPR002012; GnRH.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBERN1.
 DR PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Pyrrolidone carboxylic acid.
 FT NON_TER 1 1
 FT CHAIN 1 >61 PROGONADOLIBERIN I.
 FT PEPTIDE 1 10 GONADOLIBERIN I.
 FT PEPTIDE 14 >61 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 ACTIVITY.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).
 FT NON_TER 61 61
 SQ SEQUENCE 61 AA; 6828 MW; 63962AAAE319B8F0 CRC64;

Query Match 23.2%; Score 61.5; DB 1; Length 61;
 Best Local Similarity 42.9%; Pred. No. 0.37;

Matches 15; Conservative 3; Mismatches 2; Indels 15; Gaps 2;
 QY 2 HWSYGLRPGSSSLDEKKIAKMEKASSVFVNSSSGP
 |||||
 Db 2 HWSYGLRPGG-----KRNAK-----NVIDS 21
 |||||

RESULT 12
 GON3 SPAAU
 ID GON3 SPAAU STANDARD; PRT; 90 AA.
 AC P51923;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gonadoliberein III precursor (Gonadotropin-releasing hormone III)
 (GNRH-III) (LH-RH III) (Luliberin III).
 GN GNRH3.
 OS Sparus aurata (Gilthead sea bream).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
 OC Sparidae; Sparus.
 OX NCBI_TaxID=8175;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99061849; PubMed=9843645;
 RA Holland M.C.H., Gothelf Y., Meiri I., King J.A., Okuzawa K.,
 Eliur A., Zohar Y.;
 RT "Levels of the native forms of GnRH in the pituitary of the gilthead
 seabream, Sparus aurata, at several characteristic stages of the
 gonadal cycle.";
 RL Gen. Comp. Endocrinol. 112:394-405(1998).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnRH family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; U30311; AAA98445.1; -.
 DR InterPro; IPR002012; GnRH.
 DR Pfam; PF00446; GnRH; 1.
 DR PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Signal; Pyrrolidone carboxylic acid.
 FT SIGNAL 1 23 BY SIMILARITY.
 FT CHAIN 24 90 PROGONADOLIBERIN III.
 FT PEPTIDE 24 33 GONADOLIBERIN III.
 FT PEPTIDE 37 82 GNRH-ASSOCIATED PEPTIDE III (POTENTIAL).
 FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID
 (BY SIMILARITY).
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 90 AA; 10030 MW; FB4E47EB86C3FBD CRC64;

Query Match 23.2%; Score 61.5; DB 1; Length 90;
 Best Local Similarity 30.4%; Pred. No. 0.58;
 Matches 17; Conservative 6; Mismatches 16; Indels 17; Gaps 2;

QY 2 HWSYGLRPGKRSVGELEATRMGTGGVSLPPEASQTQERLPYNNVKDDSP 80
 |||||
 Db 25 HWSYGLPGKRSVGELEATRMGTGGVSLPPEASQTQERLPYNNVKDDSP 80
 |||||

RESULT 13
 GON3 DICLA
 ID GON3 DICLA STANDARD; PRT; 90 AA.
 AC Q91A09;

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gonadoliberin III precursor (Gonadotropin-releasing hormone III)
 DE (GNRH-III) (LH-RH III) (Luliberin III)
 GN GNRH3.
 OS Dicertrarchus labrax (European sea bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 OC Moronidae; Dicertrarchus.
 OX NCBI_TaxID=13489;
 RN SEQUENCE FROM N.A.
 RP MEDLINE=20540016; PubMed=11086295;
 RC TISSUE=Brain;
 RA Gonzalez-Martinez D., Madigou T., Zmora N., Anglade I., Zanuy S.,
 RA Zohar Y., Elizur A., Munoz-Cueto J.A., Kah O.;
 RT "Differential expression of three different prepro-GNRH
 RT (gonadotropin-releasing hormone) messengers in the brain of the
 RT European sea bass (Dicertrarchus labrax).";
 RL J. Comp. Neurol. 429:144-155(2001).
 CC -1- FUNCTION: Stimulates the secretion of gonadotropins (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the GNRH family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF224280; AAF62899.1; -
 DR InterPro: IPR002012; GNRH.
 DR Pfam; PF00446; GNRH. 1
 DR PROSITE; PS00473; GNRH. 1
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Signal; Pyrrolidone carboxylic acid.
 FT SIGNAL 1 23 BY SIMILARITY.
 FT CHAIN 24 90 PROGNADOLIBERIN III.
 FT PEPTIDE 24 33 GONADOLIBERIN III.
 FT PEPTIDE 37 82 GNRH-ASSOCIATED PEPTIDE III (POTENTIAL).
 FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID
 FT (BY SIMILARITY).
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 90 AA; 10154 MW; B06A7BA13930C67 CRC64;
 Query Match 22.8%; Score 60.5; DB 1; Length 90;
 Best Local Similarity 31.5%; Pred. No. 0.78;
 Matches 17; Conservative 6; Mismatches 14; Indels 17; Gaps 2;
 QY 2 HWSYGLRPG---SSGP-----SLDEKKIANMEKASSVFNVNSSS 38
 DB 25 HWSYGLPGKRSVGELEATIRMMGTGEVSLPEASQAQTERLPYINVDSS 78
 RESULT 14
 GONI_HAPBU
 ID GONI_HAPBU STANDARD; PRT; 94 AA.
 AC P51918; O93387;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Prognadoliberin I precursor (Contains: Gonadoliberin I (Luteinizing
 DE hormone releasing hormone I) (Gonadotropin-releasing hormone I)
 DE (GNRH-I) (LH-RH I) (Luliberin I); GNRH-associated peptide I).
 GN GNRH1.
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidel;
 OC Cichlidae; Astatotilapia.
 OX NCBI_TaxID=8133;
 RN SEQUENCE FROM N.A.
 RP MEDLINE=95396797; PubMed=7667296;
 RA White S.A., Kasten T.L., Bond C.T., Adelman J.P., Fernald R.D.;
 RT "Three gonadotropin-releasing hormone genes in one organism suggest
 RT novel roles for an ancient peptide.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:8363-8367(1995).
 RN SEQUENCE FROM N.A.
 RP MEDLINE=99061842; PubMed=9843638;
 RA White R.B., Fernald R.D.;
 RT "Ontogeny of gonadotropin-releasing hormone (GNRH) gene expression
 RT reveals a distinct origin for GNRH-containing neurons in the
 RT midbrain.";
 RL Gen. Comp. Endocrinol. 112:322-329(1998).
 RN SEQUENCE OF 23-32, AND MASS SPECTROMETRY.
 RC TISSUE=Pituitary;
 RX MEDLINE=95372591; PubMed=7644702;
 RA Powell J.F.F., Fischer W.H., Park M., Craig A.G., Rivier J.E.,
 RA White S.A., Francis R.C., Fernald R.D., Licht P., Warby C.,
 RA Sherwood N.M.;
 RA "Primary structure of solitary form of gonadotropin-releasing hormone
 RT (GNRH) in cichlid pituitary; three forms of GNRH in brain of cichlid
 RT and pumpkinseed fish.";
 RL Regul. Pept. 57:43-53(1995).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE
 CC RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALAMIC-PITUITARY-
 CC GONADAL AXIS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS
 CC TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPOTHYSEAL AXONS.
 CC -1- MASS SPECTROMETRY: MW=1113.9; METHOD=MALDI; RANGE=23-32.
 CC -1- SIMILARITY: Belongs to the GNRH family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U01865; AAC59691.1; -
 DR EMBL; AF076961; AAC27716.1; -
 DR PIR; I50739; I50739.
 DR GO; GO:0005576; C:extracellular; NAS.
 DR GO; GO:0005183; P:luteinizing hormone-releasing factor activity; NAS.
 DR GO; GO:0007275; P:development; IDA.
 DR InterPro: IPR002012; GNRH.
 DR InterPro: IPR004079; GonadoliberinI.
 DR Pfam; PF00446; GNRH; 1.
 DR PRINTS; PR01541; GONADOLIBERNI.
 DR PROSITE; PS00473; GNRH; 1.
 DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Signal; Multigene family; Pyrrolidone carboxylic acid.
 FT SIGNAL 1 22
 FT CHAIN 23 94 PROGNADOLIBERIN I.
 FT PEPTIDE 23 32 GONADOLIBERIN I.
 FT PEPTIDE 36 94 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
 FT MOD_RES 23 23 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 32 32 AMIDATION (G-33 PROVIDE AMIDE GROUP).
 FT CONFLICT 86 94 ENGHRTFKK -> KMDTGHRSNRFL (IN REF. 1).
 SQ SEQUENCE 94 AA; 10382 MW; E57DBA8333278D7 CRC64;
 Query Match 22.6%; Score 60; DB 1; Length 94;
 Best Local Similarity 30.4%; Pred. No. 0.95;
 Matches 17; Conservative 6; Mismatches 23; Indels 10; Gaps 2;
 QY 2 HWSYGLRPGSGPSLDE-----KKIAKMEKASSVFNVNSSSGPSLHWSYGL 48

Db 24 HWSYGLSPGK-RDLNFSDTLGNMVEEPFPCVFCABESPPFAXMYRVKGL 78

RESULT 15

```

THIC_SHEON
ID THIC_SHEON STANDARD; PRT; 721 AA.
AC Q8EED7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Thiamine biosynthesis protein thic.
GN THIC OR S02445.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1.
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Neilson K.H., Fraser C.M.;
RA "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123(2002).
CC -!- FUNCTION: Required for the synthesis of the hydromethylpyrimidine
CC (HMP) moiety of thiamine (4-amino-2-methyl-5-
CC hydroxymethylpyrimidine) (By similarity).
CC -!- PATHWAY: Thiamine biosynthesis.
CC -!- SIMILARITY: Belongs to the thic family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE015685; AAN55479.1; -
CC TIGR; SO2445; -
CC HAMAP; MF_00089; -; 1.
CC InterPro; IPR002817; Thic.
CC Pfam; PF01964; Thic; 1.
CC ProDom; PD007048; Thic; 1.
CC TIGRFAMs; TIGR00190; thic; 1.
CC Thiamine biosynthesis; Complete proteome.
CC SEQUENCE 721 AA; 79501 MW; C3A9B5D0F3D3D4D4 CRC64;
DR EMBL; AE015685; AAN55479.1; -
DR TIGR; SO2445; -
DR HAMAP; MF_00089; -; 1.
DR InterPro; IPR002817; Thic.
DR Pfam; PF01964; Thic; 1.
DR ProDom; PD007048; Thic; 1.
DR TIGRFAMs; TIGR00190; thic; 1.
KW Thiamine biosynthesis; Complete proteome.
SQ SEQUENCE 721 AA; 79501 MW; C3A9B5D0F3D3D4D4 CRC64;
Query Match 22.5%; Score 59.5; DB 1; Length 721;
Best Local Similarity 27.3%; Pred. NO.12;
Matches 12; Conservative 9; Mismatches 14; Indels 9; Gaps 1;
QY 6 GLRPGSGPSLDEKKIAKVEKASSVFNVNNGSGPSLHWSYGLR 49
Db 412 GWPGSIADANDEAQAELTIGELVNIA-----WEYDVQ 446

```

Search completed: March 10, 2004, 09:14:00
Job time : 10.3268 secs


```
Q25729 ID Q25729 PRELIMINARY; PRT; 408 AA.
AC Q25729;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata, Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Santa Lucia;
RA Qari S.H., Lal A.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20969; AA63153.1; -
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR00367; Crmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PRO1303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Malaria.
SQ SEQUENCE 408 AA; 43871 MW; A545BE517822515C CRC64;

Query Match 41.5%; Score 110; DB 5; Length 408;
Best Local Similarity 61.0%; Pred. No. 1e-05;
Matches 25; Conservative 3; Mismatches 5; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAMEKASSVFNVNSSG 39
DB 357 IKPGSAGSKDELVDYNDIEKKICKMEKCSSVFNVNSSIG 397

RESULT 3
Q27325 ID Q27325 PRELIMINARY; PRT; 436 AA.
AC Q27325;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata, Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=837;
RA Dame J.B., Williams J.L., McCutchan T.F., Weber J.L., Wirtz R.A.,
RA Rockmeyer W.T., Maloy W.L., Haynes J.D., Schneider I., Roberts D.,
RA Sanders G.S., Reddy P.E., Diggs C.D., Miller L.H.;
RT "Structure of the gene encoding the immunodominant surface antigen on
RT the sporozoite of the human malaria parasite Plasmodium falciparum.";
RL Science 225:593-599(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=837;
RA la Cruz V.F.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=837;
RA Jongwutives S., Tanabe K., Kanbara H.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M83164; AAA29542.1; -
DR EMBL; M83150; AAA29563.1; -
DR EMBL; M83163; AAA29576.1; -
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
```

```
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PRO1303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
SQ SEQUENCE 436 AA; 46688 MW; 5B42FP3348B8655 CRC64;

Query Match 41.5%; Score 110; DB 5; Length 436;
Best Local Similarity 61.0%; Pred. No. 1.1e-05;
Matches 25; Conservative 3; Mismatches 5; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKKIAMEKASSVFNVNSSG 39
DB 385 IKPGSAGSKDELVDYNDIEKKICKMEKCSSVFNVNSSIG 425

RESULT 4
Q819H8 ID Q819H8 PRELIMINARY; PRT; 392 AA.
AC Q819H8;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata, Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Asemo Bay;
RX MEDLINE=22356746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magalis M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
RT transmission areas-XVI: Asemo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90(2002).
DR EMBL; AF540461; AAN87595.1; -
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PRO1303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
SQ SEQUENCE 392 AA; 42385 MW; 99DAAD1629801E0C CRC64;

Query Match 40.0%; Score 106; DB 5; Length 392;
Best Local Similarity 63.4%; Pred. No. 3.3e-05;
Matches 26; Conservative 3; Mismatches 4; Indels 8; Gaps 2;

QY 7 LRFSSGSPS---LD-----EKKIAMEKASSVFNVNSSG 39
DB 341 IKPGSAGSKDELVDYNDIEKKICKMEKCSSVFNVNSSIG 381

RESULT 5
Q819I1 ID Q819I1 PRELIMINARY; PRT; 396 AA.
AC Q819I1;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata, Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Boliviar;
RX MEDLINE=22356746; PubMed=12467976;
```



```
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
RT transmission areas-XVI, Asembo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90(2002).
DR EMBL; AF540456; AAN87592.1; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
SQ SEQUENCE 396 AA; 42781 MW; 97070A9ED8D517D1 CRC64;

Query Match 40.0%; Score 106; DB 5; Length 396;
Best Local Similarity 63.4%; Pred. No. 3.3e-05;
Matches 26; Conservative 3; Mismatches 4; Indels 8; Gaps 2;

QY 7 LRPSSGSPS---LD-----EKKIAMEKASSVFNVNVSIG 39
DB 345 IKPGSAGSKNLDYENDIEKKICKMEKCSSVFNVNVSIG 385

RESULT 6
Q819H7 PRELIMINARY; PRT; 396 AA.
AC Q819H7;
RX STRAIN=Amazonas;
RX MEDLINE=22356746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
RT transmission areas-XVI, Asembo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90(2002).
DR EMBL; AF540456; AAN87592.1; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
SQ SEQUENCE 396 AA; 42782 MW; 87564F9AD8D517D1 CRC64;

Query Match 40.0%; Score 106; DB 5; Length 396;
Best Local Similarity 63.4%; Pred. No. 3.3e-05;
Matches 26; Conservative 3; Mismatches 4; Indels 8; Gaps 2;

QY 7 LRPSSGSPS---LD-----EKKIAMEKASSVFNVNVSIG 39
DB 345 IKPGSAGSKNLDYENDIEKKICKMEKCSSVFNVNVSIG 385

RESULT 7
Q819H9 PRELIMINARY; PRT; 396 AA.
AC Q819H9;
RX STRAIN=Amazonas;
RX MEDLINE=22356746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
RT transmission areas-XVI, Asembo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90(2002).
DR EMBL; AF540456; AAN87592.1; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
SQ SEQUENCE 396 AA; 42782 MW; 87564F9AD8D517D1 CRC64;

Query Match 40.0%; Score 106; DB 5; Length 396;
Best Local Similarity 63.4%; Pred. No. 3.3e-05;
Matches 26; Conservative 3; Mismatches 4; Indels 8; Gaps 2;

QY 7 LRPSSGSPS---LD-----EKKIAMEKASSVFNVNVSIG 39
DB 345 IKPGSAGSKNLDYENDIEKKICKMEKCSSVFNVNVSIG 385
```

```
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Amazonas;
RX MEDLINE=22356746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
RT transmission areas-XVI, Asembo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90(2002).
DR EMBL; AF540460; AAN87594.1; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
SQ SEQUENCE 396 AA; 42768 MW; 709FA806690FD17C CRC64;

Query Match 40.0%; Score 106; DB 5; Length 396;
Best Local Similarity 63.4%; Pred. No. 3.3e-05;
Matches 26; Conservative 3; Mismatches 4; Indels 8; Gaps 2;

QY 7 LRPSSGSPS---LD-----EKKIAMEKASSVFNVNVSIG 39
DB 345 IKPGSAGSKNLDYENDIEKKICKMEKCSSVFNVNVSIG 385

RESULT 8
Q819H7 PRELIMINARY; PRT; 360 AA.
AC Q819H7;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Asembo Bay;
RX MEDLINE=22356746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
RT transmission areas-XVI, Asembo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90(2002).
DR EMBL; AF540462; AAN87596.1; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crmsprzoite.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01303; CRMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
SQ SEQUENCE 360 AA; 39093 MW; 6504CC012649236C CRC64;

Query Match 39.6%; Score 105; DB 5; Length 360;
Best Local Similarity 58.5%; Pred. No. 4e-05;
Matches 24; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPSSGSPS---LD-----EKKIAMEKASSVFNVNVSIG 39
DB 309 IKPGSAGSKNLDYENDIEKKICKMEKCSSVFNVNVSIG 349
```

```

SQ SEQUENCE 412 AA; 44304 MW; 5D6854F31AEF949A CRC64;
Query Match 39.6%; Score 105; DB 5; Length 412;
Best Local Similarity 58.5%; Pred. No. 4.7e-05;
Matches 24; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKIKAKMEKASSVFNVNSSG 39
DB 361 IKPGSAGPKDELDYANDIEKKICKMEKCSSVFNVNSSIG 401

RESULT 11
Q25838 PRELIMINARY; PRT; 420 AA.
AC Q25838;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=835b;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutiwes S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
falciparum from Thai field isolates.";
RL Am. J. Trop. Med. Hyg. 51:659-668 (1994).
DR EMBL; M83161; AAA29574.1;
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1;
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
SQ SEQUENCE 420 AA; 45155 MW; 3A85B92432C2893C CRC64;
Query Match 39.6%; Score 105; DB 5; Length 420;
Best Local Similarity 58.5%; Pred. No. 4.8e-05;
Matches 24; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKIKAKMEKASSVFNVNSSG 39
DB 369 IKPGSAGPKDELDYANDIEKKICKMEKCSSVFNVNSSIG 409

RESULT 12
Q819J2 PRELIMINARY; PRT; 389 AA.
AC Q819J2;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Asemo Bay;
RX MEDLINE=22356746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
circumsporozoite protein (CSP) of Plasmodium falciparum from different
transmission areas-XVI. Asemo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90 (2002).
DR EMBL; AF540463; AAN87577.1;

```

```

RESULT 9
Q819H6 PRELIMINARY; PRT; 360 AA.
AC Q819H6;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Asemo Bay;
RX MEDLINE=22356746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
circumsporozoite protein (CSP) of Plasmodium falciparum from different
transmission areas-XVI. Asemo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90 (2002).
DR EMBL; AF540463; AAN87577.1;
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
SQ SEQUENCE 360 AA; 39062 MW; 65059844270D666C CRC64;
Query Match 39.6%; Score 105; DB 5; Length 360;
Best Local Similarity 58.5%; Pred. No. 4e-05;
Matches 24; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPSSGSPSLD-----EKIKAKMEKASSVFNVNSSG 39
DB 309 IKPGSAGPKDELDYANDIEKKICKMEKCSSVFNVNSSIG 349

RESULT 10
Q819H5 PRELIMINARY; PRT; 412 AA.
AC Q819H5;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Amazonas;
RX MEDLINE=22356746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Isea R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
circumsporozoite protein (CSP) of Plasmodium falciparum from different
transmission areas-XVI. Asemo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90 (2002).
DR EMBL; AF540463; AAN87598.1;
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro; IPR003067; Crcmsprzoite.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01303; CRCMSPRZOITE.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.

```

```

DR GO: GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro: IPR003067; Crmsprzoite.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; Crmsprzoite.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
DR SEQUENCE 389 AA; 41980 MW; 4P2C32A159B13F20 CRC64;

Query Match 38.5%; Score 102; DB 5; Length 393;
Best Local Similarity 61.0%; Pred. No. 0.00011;
Matches 25; Conservative 3; Mismatches 5; Indels 8; Gaps 2;

QY 7 LRPSSGP---SLD-----EKKIAKMEKASSVFVNVSIG 39
:||||:||||:||||:||||:||||:||||:||||:||||:
DB 338 IKPGSAGKPKNELDYENDIEKKICKMEKCSSVFVNVSIG 378

RESULT 13
Q819J1 PRELIMINARY; PRT; 393 AA.
AC Q819J1
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Yacoude;
RX MEDLINE=22355746; PubMed=12467976;
RA Escalante A.A., Grebert H.M., Issa R., Goldman I.F., Basco L.,
RA Magris M., Biswas S., Kariuki S., Lal A.A.;
RT "A study of genetic diversity in the gene encoding the
RT circumsporozoite protein (CSP) of Plasmodium falciparum from different
RT transmission areas-XVI. Asempo Bay Cohort Project.";
RL Mol. Biochem. Parasitol. 125:83-90(2002).
RL EMBL; AF540444; AAN87578.1; -.
DR GO: GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro: IPR003067; Crmsprzoite.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; Crmsprzoite.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
DR SEQUENCE 393 AA; 42276 MW; 7D2C7CCFAA3153A9 CRC64;

Query Match 38.5%; Score 102; DB 5; Length 393;
Best Local Similarity 61.0%; Pred. No. 0.00011;
Matches 25; Conservative 3; Mismatches 5; Indels 8; Gaps 2;

QY 7 LRPSSGP---SLD-----EKKIAKMEKASSVFVNVSIG 39
:||||:||||:||||:||||:||||:||||:||||:||||:
DB 342 IKPGSAGKPKNELDYENDIEKKICKMEKCSSVFVNVSIG 382

RESULT 14
Q8MZK7 PRELIMINARY; PRT; 69 AA.
AC Q8MZK7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Circumsporozoite protein II (fragment).
GN CSPII.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=PCCI/HN;
RA Chen H.H., Yu X.B., Wu Z.D., Xu J.;
RT "Gene cloning of circumsporozoite protein (CSP) II gene from
RT Plasmodium falciparum (PCCI/HN).";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY093672; AAM19072.1; -.
DR GO: GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro: IPR003067; Crmsprzoite.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; Crmsprzoite.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
DR NON_TER 1
FT SEQUENCE 69 AA; 7670 MW; F3ELAA54671FE6FF CRC64;

Query Match 37.4%; Score 99; DB 5; Length 69;
Best Local Similarity 56.1%; Pred. No. 3.2e-05;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGP---SLD-----EKKIAKMEKASSVFVNVSIG 39
:||||:||||:||||:||||:||||:||||:||||:||||:
DB 18 IKPGSANKPKQLDYENDIEKKICKMEKCSSVFVNVSIG 58

RESULT 15
Q25839 PRELIMINARY; PRT; 115 AA.
AC Q25839;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Circumsporozoite protein (fragment).
GN CSP.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=835c;
RX MEDLINE=95077069; PubMed=7985759;
RA Jongwutives S., Tanabe K., Hughes M.K., Kanbara H., Hughes A.L.;
RT "Allelic variation in the circumsporozoite protein of Plasmodium
RT falciparum from Thai field isolates";
RL Am. J. Trop. Med. Hyg. 51:659-668(1994).
DR EMBL; M83162; AAA29575.1; -.
DR GO: GO:0003793; F:defense/immunity protein activity; IEA.
DR InterPro: IPR003067; Crmsprzoite.
DR Pfam: PF00090; tsp_1; 1.
DR PRINTS: PR01303; Crmsprzoite.
DR SMART: SM00209; TSP1; 1.
DR PROSITE: PS50092; TSP1; 1.
DR NON_TER 1
FT SEQUENCE 115 AA; 12974 MW; 6BB538287260DA90 CRC64;

Query Match 37.4%; Score 99; DB 5; Length 115;
Best Local Similarity 56.1%; Pred. No. 6e-05;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRPSSGP---SLD-----EKKIAKMEKASSVFVNVSIG 39
:||||:||||:||||:||||:||||:||||:||||:||||:
DB 64 IKPGSANKPKQLDYENDIEKKICKMEKCSSVFVNVSIG 104

Search completed: March 10, 2004, 09:25:41
Job time : 50.4047 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:48 ; Search time 76.7977 Seconds
(without alignments)
187.635 Million cell updates/sec

Title: US-09-848-834A-20
Perfect score: 265
Sequence: 1 XHWSYGLRPGSSGSLDEKX.....NVNSSSGSLHWSYGLRFX 51

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1980s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	263	99.2	51	AAU11431	AAU11431 Synthetic
2	182	68.7	36	AAU11427	AAU11427 Synthetic
3	160.5	60.6	46	AAU11430	AAU11430 Synthetic
4	157	59.2	33	AAU11423	AAU11423 Synthetic
5	144	54.3	47	AAU11428	AAU11428 Synthetic
6	138.5	52.3	50	AAU11429	AAU11429 Synthetic
7	127.5	48.1	33	AAR62715	AAR62715 LHRH-cont
8	99	37.4	309	AAR13175	AAR13175 NSI_81-RL
9	99	37.4	319	AAR07945	AAR07945 NSI81RLFA
10	99	37.4	319	AAR11176	AAR11176 NSI_81-RL
11	99	37.4	327	AAR13177	AAR13177 NSI_81-RL
12	99	37.4	335	AAR13178	AAR13178 NSI_81(NV
13	99	37.4	335	AAR13179	AAR13179 NSI_81(NV
14	99	37.4	396	ABC23530	ABC23530 Plasmodiu
15	99	37.4	411	AAP83144	AAP83144 Sequence
16	99	37.4	412	AAP60416	AAP60416 CS protei
17	99	37.4	424	AAR37797	AAR37797 RTS+ prot
18	97.5	36.8	49	AAW03944	AAW03944 GnRH 4-re
19	97.5	36.8	49	AAW79567	AAW79567 GnRH-2. 1
20	97.5	36.8	49	AAW61542	AAW61542 Peptide h
21	97.5	36.8	49	AAV58363	AAV58363 Four-copy
22	97.5	36.8	49	AAV58135	AAV58135 GnRH anal
23	97.5	36.8	544	AAW03943	AAW03943 LKT-GnRH
24	97.5	36.8	544	AAW79570	AAW79570 LKT-GnRH
25	97.5	36.8	695	AAW79573	AAW79573 LKT-GnRH

26	97.5	36.8	695	3	AAV58361	AAV58361 Leukotoxi
27	97.5	36.8	695	3	AAV58133	AAV58133 Gonadotro
28	97.5	36.8	977	2	AAW03942	AAW03942 LKT-GnRH
29	97.5	36.8	977	2	AAW79569	AAW79569 LKT-GnRH
30	95	35.8	20	5	AAU11414	AAU11414 P. falcip
31	95	35.8	350	3	AAV70278	AAV70278 Reconbina
32	95	35.8	412	1	AAW80835	AAW80835 Sequence
33	94	35.5	143	3	AAV49252	AAV49252 N6 polyep
34	94	35.5	218	3	AAV49253	AAV49253 N10 polye
35	94	35.5	240	3	AAV49254	AAV49254 N11 polye
36	94	35.5	390	3	AAV49255	AAV49255 N19 polye
37	91	34.3	424	2	AAR37796	AAR37796 RTS prote
38	90.5	34.2	42	3	AAW20865	AAW20865 GnRH tand
39	90	34.0	23	3	AAW20864	AAW20864 GnRH tand
40	89.5	33.8	33	4	AAW63463	AAW63463 Peptide c
41	89.5	33.8	33	4	AAW63463	AAW63463 Peptide c
42	89	33.6	19	4	AAW98951	AAW98951 Vaccine r
43	89	33.6	21	1	AAW91504	AAW91504 Sequence
44	89	33.6	21	2	AAW78920	AAW78920 Malaria c
45	89	33.6	21	2	AAW75955	AAW75955 P. falcip

ALIGNMENTS

RESULT 1
AAU11431
ID AAU11431 standard; peptide; 51 AA.
XX
AC AAU11431;
XX
DT 12-MAR-2002 (first entry)
XX
DE Synthetic immunogen peptide 12.
XX
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS Plasmodium falciparum.
OS Mammalia.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Peptide 1..10
/note= "Gonadotrophin releasing hormone epitope (1..10 aa)"
FT Misc-difference 1
/label= OTHER
FT Peptide /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT Peptide 11..16
/note= "Spacer peptide"
FT Peptide 17..36
/note= "Malaria CSP protein (378-398 aa)"
FT Peptide 37..42
/note= "Spacer peptide"
FT Peptide 43..51
/note= "Gonadotrophin releasing hormone epitope (2-10 aa)"
FT Modified-site 51
/note= "Amidated glycine or glycineamide"
XX
WO200185763-A2.
XX
PD 15-NOV-2001.
XX
PF 04-MAY-2001; 2001WC-US014363.
XX
PR 05-MAY-2000; 2000US-0202328P.
XX

PA (APHT-) APHTON CORP.
 XX Grimes S, Michaeli D, Stevens VC;
 PI WPI; 2002-049440/06.
 XX Novel synthetic immunogen for inducing immune response against
 XX gonadotropin releasing hormone, comprises fusion peptide having
 XX promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 XX or its analog.
 XX Claim 11; Page 12-13; 43pp; English.
 XX The invention relates to a synthetic immunogen for inducing specific
 XX antibodies against gonadotropin releasing hormone (GnRH) also known as
 XX luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
 XX which comprises a promiscuous helper T-cell peptide epitope and
 XX immunomimic peptide epitope or its analogue. The synthetic immunogen is
 XX useful inducing an immune response against GnRH in an animal subject, and
 XX as such is useful as a contraceptive and in the treatment of diseases
 XX such as cancer (of the breast, uterus and other gynaecological cancer),
 XX endometriosis, uterine fibroids, benign prostatic hypertrophy and
 XX prostate cancer. The immunogen is effective in eliciting high and
 XX specific anti-GnRH antibody titres. The present sequence is a synthetic
 XX immunogen of the invention
 XX Sequence 51 AA;
 SQ Query Match 99.2%; Score 263; DB 5; Length 51;
 Best Local Similarity 100.0%; Pred. No. 9.4e-27;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSGPSLDEKKIAMEKASVFNVNSSGSPSLHWSYGLRP 50
 DB 2 HWSYGLRPGSGPSLDEKKIAMEKASVFNVNSSGSPSLHWSYGLRP 50

RESULT 2
 AAU11427
 ID AAU11427 standard; peptide; 36 AA.
 AC AAU11427;
 XX
 DT 12-MAR-2002 (first entry)
 DE Synthetic immunogen peptide 8.
 XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
 KW luteinising hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
 XX Plasmodium falciparum.
 OS Mammalia.
 OS Synthetic.
 OS Chimeric.
 XX Key Location/Qualifiers
 FH Peptide 1..10
 FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
 FT /note= "Spacer peptide"
 FT /note= "Spacer peptide"
 FT /note= "Malaria CSP protein (378-398 aa)"
 XX WC200185763-A2.
 PN 15-NOV-2001.
 XX

PF 04-MAY-2001; 2001WO-US014363.
 XX
 PR 05-MAY-2000; 2000US-0202328P.
 XX (APHT-) APHTON CORP.
 XX Grimes S, Michaeli D, Stevens VC;
 PI WPI; 2002-049440/06.
 XX Novel synthetic immunogen for inducing immune response against
 XX gonadotropin releasing hormone, comprises fusion peptide having
 XX promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 XX or its analog.
 XX Claim 11; Page 10; 43pp; English.
 XX The invention relates to a synthetic immunogen for inducing specific
 XX antibodies against gonadotropin releasing hormone (GnRH) also known as
 XX luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
 XX which comprises a promiscuous helper T-cell peptide epitope and
 XX immunomimic peptide epitope or its analogue. The synthetic immunogen is
 XX useful inducing an immune response against GnRH in an animal subject, and
 XX as such is useful as a contraceptive and in the treatment of diseases
 XX such as cancer (of the breast, uterus and other gynaecological cancer),
 XX endometriosis, uterine fibroids, benign prostatic hypertrophy and
 XX prostate cancer. The immunogen is effective in eliciting high and
 XX specific anti-GnRH antibody titres. The present sequence is a synthetic
 XX immunogen of the invention
 XX Sequence 36 AA;
 SQ Query Match 68.7%; Score 182; DB 5; Length 36;
 Best Local Similarity 100.0%; Pred. No. 2.3e-16;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSGPSLDEKKIAMEKASVFNVNSS 36
 DB 2 HWSYGLRPGSGPSLDEKKIAMEKASVFNVNSS 36

RESULT 3
 AAU11430
 ID AAU11430 standard; peptide; 46 AA.
 AC AAU11430;
 XX
 DT 12-MAR-2002 (first entry)
 DE Synthetic immunogen peptide 11.
 XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
 KW luteinising hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
 XX Clostridium tetani.
 OS Mammalia.
 OS Synthetic.
 OS Chimeric.
 XX Key Location/Qualifiers
 FH Peptide 1..10
 FT /note= "Gonadotropin releasing hormone epitope (1..10 aa)"
 FT Misc-difference 1
 FT /label= OTHER
 FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
 FT /note= "Spacer peptide"
 FT /note= "Spacer peptide"
 FT /note= "Tetanus toxoid (830-844 aa)"
 FT

FT Peptide 32. .37
 FT /note= "Spacer peptide"
 FT Peptide 38. .46
 FT /note= "Gonadotropin releasing hormone epitope (2-10 aa)"
 FT Modified-site 46
 FT /note= "Amidated glycine or glycylamide"
 XX WO200185763-A2.
 XX 15-NOV-2001.
 XX 04-MAY-2001; 2001WO-US014363.
 XX 05-MAY-2000; 2000US-0202328P.
 XX (APHT-) APHTON CORP.
 XX Grimes S, Michaeli D, Stevens VC;
 XX WPI; 2002-049440/06.
 XX Novel synthetic immunogen for inducing immune response against
 FT gonadotropin releasing hormone, comprises fusion peptide having
 FT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 FT or its analog.
 XX Claim 11; Page 12; 43pp; English.
 XX The invention relates to a synthetic immunogen for inducing specific
 CC antibodies against gonadotropin releasing hormone (GnRH) also known as
 CC luteinising hormone releasing hormone (LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention
 XX Sequence 46 AA;
 SQ
 Query Match 60.6%; Score 160.5; DB 5; Length 46;
 Best Local Similarity 66.0%; Pred. No. 2e-13;
 Matches 33; Conservative 3; Mismatches 7; Indels 7; Gaps 2;
 QY 2 HWSYGLRPGSGPSLDEKKIAKMEKASSVF-NVNVSSGSPSLHWSYGLRP 50
 DB 2 HWSYGLRPGSGPSL-----QYIKANSKFIGITELSSGSPSLHWSYGLRP 45
 RESULT 4
 AAU11423
 ID AAU11423 standard; peptide; 33 AA.
 XX
 AC AAU11423;
 XX
 XX 12-MAR-2002 (first entry)
 XX
 DE Synthetic immunogen peptide 4.
 XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
 KW luteinising hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
 XX Plasmodium falciparum.
 OS Mammalia.
 OS Synthetic.
 OS Chimeric.

XX Key Location/Qualifiers
 FH Peptide 1. .20
 FT /note= "Malaria CSP protein (378-398 aa)"
 FT Peptide 21. .24
 FT /note= "Spacer peptide"
 FT Peptide 25. .33
 FT /note= "Gonadotropin releasing hormone epitope"
 FT Modified-site 33
 FT /note= "Amidated glycine or glycylamide"
 XX WO200185763-A2.
 XX 15-NOV-2001.
 XX 04-MAY-2001; 2001WO-US014363.
 XX 05-MAY-2000; 2000US-0202328P.
 XX (APHT-) APHTON CORP.
 XX Grimes S, Michaeli D, Stevens VC;
 XX WPI; 2002-049440/06.
 XX Novel synthetic immunogen for inducing immune response against
 FT gonadotropin releasing hormone, comprises fusion peptide having
 FT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 FT or its analog.
 XX Claim 11; Page 8; 43pp; English.
 XX The invention relates to a synthetic immunogen for inducing specific
 CC antibodies against gonadotropin releasing hormone (GnRH) also known as
 CC luteinising hormone releasing hormone (LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention
 XX Sequence 33 AA;
 SQ
 Query Match 59.2%; Score 157; DB 5; Length 33;
 Best Local Similarity 94.1%; Pred. No. 3.8e-13;
 Matches 32; Conservative 0; Mismatches 0; Indels 2; Gaps 1;
 QY 17 DEKKIAKMEKASSVFNVNVSSGSPSLHWSYGLRP 50
 DB 1 DEKKIAKMEKASSVFNVN--SGPSLHWSYGLRP 32
 RESULT 5
 AAU11428
 ID AAU11428 standard; peptide; 47 AA.
 XX
 AC AAU11428;
 XX
 XX 12-MAR-2002 (first entry)
 XX
 DE Synthetic immunogen peptide 9.
 XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
 KW luteinising hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
 XX Plasmodium falciparum.
 OS

XX	DT	12-MAR-2002 (first entry)
XX	DE	Synthetic immunogen peptide 10.
XX	XX	Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
XX	XX	luteinising hormone releasing hormone; LHRH; contraceptive;
XX	XX	promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
XX	XX	breast cancer; uterine cancer; gynaecological cancer; endometriosis;
XX	XX	uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX	XX	Clostridium tetani.
XX	XX	Mammalia.
XX	XX	Synthetic.
XX	XX	Chimeric.
XX	XX	Key
XX	XX	Location/Qualifiers
XX	XX	1..10
XX	XX	/note= "Gonadotrophin releasing hormone epitope (1..10 aa)"
XX	XX	Misc-difference 1
XX	XX	/label= OTHER
XX	XX	/note= "Other= Pyro-glutamic acid or 5-oxo proline"
XX	XX	11..16
XX	XX	/note= "Spacer peptide"
XX	XX	17..37
XX	XX	/note= "Tetanus toxoid (947-967 aa)"
XX	XX	38..41
XX	XX	/note= "Spacer peptide"
XX	XX	42..50
XX	XX	/note= "Gonadotrophin releasing hormone epitope (2-10 aa)"
XX	XX	Modified-site 50
XX	XX	/note= "Amidated glycine or glycineamide"
XX	XX	WO200185763-A2.
XX	XX	15-NOV-2001.
XX	XX	04-MAY-2001; 2001WO-US014363.
XX	XX	05-MAY-2000; 2000US-0202328P.
XX	XX	(APHT-) APHTON CORP.
XX	XX	Grimes S, Michaeli D, Stevens VC;
XX	XX	WPI; 2002-049440/06.
XX	XX	Novel synthetic immunogen for inducing immune response against
XX	XX	gonadotrophin releasing hormone, comprises fusion peptide having
XX	XX	promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
XX	XX	or its analog.
XX	XX	Claim 11; Page 11; 43pp; English.
XX	XX	The invention relates to a synthetic immunogen for inducing specific
XX	XX	antibodies against gonadotrophin releasing hormone (GnRH) also known as
XX	XX	luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
XX	XX	which comprises a promiscuous helper T-cell peptide epitope and
XX	XX	immunomimic peptide epitope or its analogue. The synthetic immunogen is
XX	XX	useful inducing an immune response against GnRH in an animal subject, and
XX	XX	as such is useful as a contraceptive and in the treatment of diseases
XX	XX	such as cancer (of the breast, uterus and other gynaecological cancer),
XX	XX	endometriosis, uterine fibroids, benign prostatic hypertrophy and
XX	XX	prostate cancer. The immunogen is effective in eliciting high and
XX	XX	specific anti-GnRH antibody titres. The present sequence is a synthetic
XX	XX	immunogen of the invention
XX	XX	Sequence 47 AA;
XX	XX	Query Match
XX	XX	Best Local Similarity 54.3%; Score 144; DB 5; Length 47;
XX	XX	Matches 28; Conservative 8; Mismatches 9; Indels 4; Gaps 2;
XX	XX	2 HWSYGLRPGSGSLDEKXIAKMEKASSFNVNVSNGSGPSLHWSYGLRP 50
XX	XX	2 HWSYGLRPGSGSGPSL--KLLSEIK--GVIVHRLGEGVSGPSLHWSYGLRP 46
XX	XX	RESULT 6
XX	XX	AAU11429
XX	XX	ID AAU11429 standard; peptide; 50 AA.
XX	XX	AC AAU11429;

Matches 29; Conservative 4; Mismatches 11; Indels 9; Gaps 2;
QY 2 HWSYGLRPGSSGSLDEKXIA---KWEKASVFNVNSSGSLHWSYGLRP 50
DB 2 HWSYGLRPGSSGSLFNFTVTFWLRVKVSA-----SHLEGPSSLHWSYGLRP 49

RESULT 7
ID AAR62715 standard; peptide; 33 AA.
XX AAR62715;
XX AC AAR62715;
XX 25-MAR-2003 (revised)
DT 10-SEP-1995 (first entry)
XX LHRH-containing immunogenic peptide.
XX Helper T cell epitope; universal immune stimulator; invasive; haptens;
KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
KW androgen-dependent carcinoma; antitumor; infertility;
KW Plasmodium falciparum circumsporozoite.
XX Synthetic.
OS
FH Key Location/Qualifiers
FT Domain 1..21 /note= "plasmodium falciparum circumsporozoite helper T
FT cell epitope"
FT Domain 24..33
FT /note= "LHRH hapten"
FT
FT
PN WO9425060-A1.
XX 10-NOV-1994.
XX 28-APR-1994; 94WO-US004832.
XX 27-APR-1993; 93US-00057166.
PR 14-APR-1994; 94US-00229275.
XX (LADD/) LADD A E.
PA (WANG/) WANG C Y.
PA (ZAMB/) ZAMB T.
XX Ladd AE, Wang CY, Zamb T;
XX WPI; 1994-357910/44.
DR
XX Immunogenic luteinising hormone releasing hormone peptide(s) - that
PT suppress LHRH activity in males and females.
XX
PS Claim 8; Page 86; 213pp; English.
XX
CC Synthetic immunogenic peptides are provided in which a universal immune
CC stimulator is linked to a peptide or protein hapten containing B cell
CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
CC potent immune responses to the coupled peptide or protein. The stimulator
CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
CC immune response to the coupled peptide in members of a heterogeneous
CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
CC sequence from the invasive protein of Yersinia. Spacer amino acid
CC sequences (e.g. Gly-Gly) can be provided between the invasin and Th
CC domains and between the immune stimulator and hapten components. When the
CC haptens is LHRH, then optionally the invasin domain can be omitted from
CC the immune stimulator component. The present sequence represents an LHRH-
CC containing, invasin-free immunogenic peptide as above which can be used
CC as a potent vaccine for treating e.g. prostatic hyperplasia, androgen-
CC dependent carcinoma, prostatic carcinoma, testicular carcinoma, ovarian
CC endometriosis, benign uterine tumours, recurrent functional ovarian
CC cysts, (severe) premenstrual syndrome or oestrogen-dependent breast
CC cancer, or for induction of infertility. (Updated on 25-MAR-2003 to
CC correct PN field.)

XX Sequence 33 AA;
SQ
Query Match 48.1%; Score 127.5; DB 2; Length 33;
Best Local Similarity 81.8%; Pred No. 2.7e-09;
Matches 27; Conservative 0; Mismatches 3; Indels 3; Gaps 1;
QY 18 EKIKAKMEKASVFNVNSSGSLHWSYGLRP 50
DB 3 EKIKAKMEKASVFNVNSSGSLHWSYGLRP 32

RESULT 8
AAR13175 standard; protein; 309 AA.
XX AAR13175;
XX AC AAR13175;
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 29-AUG-1991 (first entry)
XX NS1_81-RLfdelta9.
XX Immunogenic determinant; circumsporozoite; CS; vaccine; malaria; hybrid;
KW influenza virus; non-structural protein 1; fusion.
XX Plasmodium falciparum.
OS Influenza virus; (A/PR/8/34/).
XX
FH Key Location/Qualifiers
FT Region 1..81 /label= N-terminal of NS1
FT /note= "Influenza virus nonstructural protein 1"
FT Peptide 82..87 /label= synthetic linker
FT Region 88 /label= artifact
FT /note= "see comments"
FT Region 89..193 /label= AAs 19-123 of CS protein
FT /note= "Region I contg. flanking region less signal
FT sequence"
FT Region 194..309 /label= AAs 297-412 of CS protein
FT /note= "Region II flanking region minus 9 N-term- inal
FT AAs"
XX EP432965-A.
XX 19-JUN-1991.
XX 06-DEC-1990; 90EP-00313257.
XX 08-DEC-1989; 89US-00447746.
XX (SMIK) SMITHKLINE BEECHAM.
PA (USSA) US SEC OF ARMY.
PA (BIOM-) BIOMEDICAL RES INST.
PA (GROS/) GROSS M S.
XX Gross MS, Gordon DM, Hollingdal MR;
PI WPI; 1991-179771/25.
XX Polypeptide comprising immunogenic determinants from P falciparum - for
DR vaccine against malaria infection in humans.
XX Example 1; Page 7; 18pp; English.
XX The polypeptide is prepd. by genetic engineering of genes encoding the P.
CC falciparum circumsporozoite (CS) protein [Dane et al., Science 225 : 593
CC (1984)], and the influenza virus non-structural protein 1 (NS1), [Baez et

[illegible]

RESULT 10	
AAR13176	
ID	AAR13176 standard; protein; 319 AA.
XX	
AC	AAR13176;
XX	
DT	24-OCT-2003 (revised)
XX	
DT	25-MAR-2003 (revised)
DT	29-AUG-1991 (first entry)
XX	
XX	
DE	NS1_81-ELFauth.

XX	
XX	Immunogenic determinant; circumsporozoite; CS; vaccine; malaria; hybrid;
KW	influenza virus; non-structural protein 1; fusion.
XX	
XX	Plasmodium falciparum.
OS	Influenza virus; (A/PR/8/34).
OS	

FT	Peptide	/note= "Influenza virus nonstructural protein 1"
FT		82. .87
FT	Region	/label= synthetic linker
FT		88
FT		/label= artifact
FT		/note= "see comments"
FT	Region	89. .193
FT		/label= AAs 19-123 of CS protein
FT		/note= "Region I contg. flanking region less signal
FT		sequence"
FT	Region	194
FT		/label= artifact
FT		/note= "see comments"
FT	Region	195. .319
FT		/label= AAs 288-412 of CS protein
FT		/note= "Region II flanking region"
FT		

```

FT      Region
PT      195. .319
TT      /label= AAS 288-412 of CS protein
FT      /note= "Region II flanking region"
XX
XX
XX      PN
XX      EP432965-A.
XX      XX
ED      19-JUN-1991.
XX
XX      XX
PF      06-DEC-1990;   90EP-00313257.
XX      XX
XX      08-DEC-1989;   89US-00447746.
XX      XX
XX      (SMIK ) SMITHKLINE BEECHAM.
PA      (USSA ) US SEC OF ARMY.
PA      (BIOM-) BIOMEDICAL RES INST.
PA      (GROS/) GROSS M S.
XX      XX
XX      Gross MS, Gordon DM, Hollingdal MR;
PI      WPI; 1991-179771/25.
XX      DR

```

XX Polypeptide comprising immunogenic determinants from P falciparum - for
PT vaccine against malaria infection in humans.
XX
XX
PS Example 2; Page 10; 18pp; English.
XX
XX The polypeptide is prepd. by genetic engineering of genes encoding the P.
CC

CC falciparum circumsporozoite (CS) protein [Dame et al., Science 225 : 593
CC (1984)], and the influenza virus non-structural protein 1 (NS1), [Baez et
CC al., Nucleic Acids Research, 8 : 5845 (1980)]. The DNA encoding the 1st
CC 81 AAs of the N-terminal of NS1 (NS1_81) is linked via a synthetic
CC sequence to DNA encoding Region I contg. flanking regionless the 18 AA
CC signal region, which in turn is fused to DNA encoding Region II-contg.
CC flanking region. This CS fusion is designated RLFAuth. The Pro residue
CC separating the Asp (at the C-terminal of the linker) from RLFAuth is an
CC arti- fact of a filled in BamHI site; the Gly separating Region I and
CC Region II-contg. CS flanking regions is an artifact of a synthetic
CC FokI/NotI linker. The peptide can be used in a vaccine for protection
CC against malaria. The complete nucleotide and AA sequences are given in EP
CC -304720, filed May 1, 1990. See also AAR12306-R12311 and AAR13175-R13179.
CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 319 AA;

Query Match 37.4%; Score 99; DB 2; Length 319;
Best Local Similarity 56.1%; Pred. No. 0.00023;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;
QY 7 LRPSSGSPSLD-----EKKIARVKEKASSVFVNSSSG 39
DQ 268 IKPGSANKPKDLDYENDIEKKICKMKCKSSVFVNSSIG 308
:||||:|||||

RESULT 11
ID AAR13177 standard; protein; 327 AA.
XX AC AAR13177;
XX 24-OCT-2003 (revised)
XX 25-MAR-2003 (revised)
XX 29-AUG-1991 (first entry)
XX NS1_81-RLFAuth + (NANP)2.
XX Immunogenic determinant; circumsporozoite; CS; vaccine; malaria; hybrid;
XX influenza virus; non-structural protein 1; fusion.
XX Plasmodium falciparum.
XX Influenza virus; (A/PR/8/34/).

XX Key Location/Qualifiers
XX Region 1..81
XX /label= N-terminal of NS1
XX /note= "Influenza virus nonstructural protein 1"
XX Peptide 92..97
XX /label= synthetic linker
XX Region 88
XX /label= artifact
XX /note= "see comments"
XX Region 89..193
XX /label= AAs 19-123 of CS protein
XX /note= "Region I contg. flanking region less signal
XX sequence"
XX Region 194..201
XX /label= immunodominant repeat region
XX /note= "two tetrapeptide repeat units"
XX Region 202
XX /label= artifact
XX /note= "see comments"
XX Region 203..327
XX /label= AAs 298-412 of CS protein
XX /note= "Region II flanking region"
XX EP432965-A.
XX 19-JUN-1991.
XX 06-DEC-1990; 90EP-00313257.

XX 08-DEC-1989; 89US-00447746.
XX (SMIK) SMITHKLINE BEECHAM.
XX (USSA) US SEC OF ARMY.
XX (BIOM-) BIOMEDICAL RES INST.
XX (GROSS/) GROSS M S.
XX Gross MS, Gordon DM, Hollingdal MR;
XX WPI; 1991-179771/25.
XX Polypeptide comprising immunogenic determinants from P falciparum - for
XX vaccine against malaria infection in humans.
XX Example 3; Page 10; 18pp; English.
XX The polypeptide is prep'd. by genetic engineering of genes encoding the P.
XX falciparum circumsporozoite (CS) protein [Dame et al., Science 225 : 593
XX (1984)], and the influenza virus non-structural protein 1 (NS1), [Baez et
XX al., Nucleic Acids Research, 8 : 5845 (1980)]. The DNA encoding the 1st
XX 81 AAs of the N-terminal of NS1 (NS1_81) is linked via a synthetic
XX sequence to DNA encoding Region I contg. flanking region less the 18 AA
XX signal region. This is linked to a synthetic sequence encoding two repeat
XX units from the immunodominant region, which in turn is fused to DNA
XX encoding Region II-contg. flanking region. The Pro residue separating the
XX Asp (at the C-terminal of the linker) from the Region I-contg. CS
XX flanking region is an artifact of a filled-in BamHI site; the Gly
XX separating the repeat units and the Region II-contg. CS flanking region
XX is an artifact of a synthetic FokI/NotI linker. The peptide can be
XX used in a vaccine for protection against malaria. See also AAR12306-
XX R12311 and AAR13175-R13179. (Updated on 25-MAR-2003 to correct PA field.)
XX (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 327 AA;

Query Match 37.4%; Score 99; DB 2; Length 327;
Best Local Similarity 56.1%; Pred. No. 0.00024;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;
QY 7 LRPSSGSPSLD-----EKKIARVKEKASSVFVNSSSG 39
DQ 276 IKPGSANKPKDLDYENDIEKKICKMKCKSSVFVNSSIG 316
:||||:|||||

RESULT 12
AAR13178
ID AAR13178 standard; protein; 335 AA.
XX AC AAR13178;
XX 24-OCT-2003 (revised)
XX 25-MAR-2003 (revised)
XX 29-AUG-1991 (first entry)
XX NS1_81 (NANP)4RLFAuth.
XX Immunogenic determinant; circumsporozoite; CS; vaccine; malaria; hybrid;
XX influenza virus; non-structural protein 1; fusion.
XX Plasmodium falciparum.
XX Influenza virus; (A/PR/8/34/).
XX Key Location/Qualifiers
XX Region 1..81
XX /label= N-terminal of NS1
XX /note= "Influenza virus nonstructural protein 1"
XX Region 82..97
XX /label= immunodominant repeat region
XX /note= "four tetrapeptide repeat units"
XX Peptide 98..103
XX /label= synthetic linker
XX Region 104

Best Local Similarity 56.1%; Pred. No. 0.00025;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;
QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNSSG 39
: : : : :
Db 284 IKPGSANKPKDLDYNDIEKKICKMEKCSSVFNVNSSIG 324

RESULT 14
ABO23530
ID ABO23530 standard; protein; 396 AA.
XX ABO23530;
XX 04-SEP-2003 (first entry)
XX Plasmodium falciparum outlier protein #2.
XX Candidate protein identification; pathogen; anti-infective;
KW outlier protein; virulence protein; antigen; drug target protein;
KW pathogenic organism; antimicrobial.
XX Plasmodium falciparum.

OS US2003039963-A1.
XX US2003039963-A1.
XX 27-FEB-2003.
XX 30-MAR-2001; 2001US-00820843.
XX 30-MAR-2001; 2001US-00820843.
XX (BRAH/) BRAHMACHARI S K.
PA (RAMA/) RAMACHANDRAN S.
PA (NAND/) NANDI T.
PA (BHIM/) BHIMARAO C.
XX Brahmachari SK, Ramachandran S, Nandi T, Bhimarao C;
PI WPI; 2003-492159/46.
XX Identifying candidate proteins useful as anti-infectives involves
PT matching outlier protein sequences with protein sequences in databases.

XX Example 7; Page 44-45; 117pp; English.
XX The present invention relates to a method for identifying candidate
CC proteins in pathogens useful as anti-infectives. The invention discloses
CC a computational method which involves the calculation of several sequence
CC attributes and their subsequent analysis results in the identification
CC of outlier proteins in different pathogens. The method is useful for the
CC identification of outlier proteins (e.g. virulence proteins, antigens or
CC proteins used as drug targets) in pathogenic organisms. The method of the
CC invention provides reproducible results as it does not depend on the
CC variable biochemical characterisation of proteins. ABO23500-ABO23617
CC represent outlier proteins identified from different pathogenic organisms
XX Sequence 396 AA;
SQ

Query Match 37.4%; Score 99; DB 7; Length 396;
Best Local Similarity 56.1%; Pred. No. 0.00031;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;
QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNSSG 39
: : : : :
Db 345 IKPGSANKPKDLDYNDIEKKICKMEKCSSVFNVNSSIG 385

RESULT 15
AAP83144
ID AAP83144 standard; protein; 411 AA.
XX AAP83144;
AC

XX 25-MAR-2003 (revised)
DT 20-NOV-1990 (first entry)
XX Sequence encoded by the circumsporozoite (CS) gene from Plasmodium
DE falciparum.
DE Vaccine; antigen; immunogen; probe; hybridisation; immunoassay;
XX KW diagnosis.
XX Plasmodium falciparum.
XX Key Location/Qualifiers
FT Region 106..120
FT /note="Region 1"
FT Region 123..146
FT /note="Repeat region, repeat unit = NANPNVDP"
FT Region 147..206
FT /note="Repeat region, repeat unit = NANP"
FT Region 211..286
FT /note="Repeat region, repeat unit = NANP"
XX EP278940-A.
XX 17-AUG-1988.
XX 25-JAN-1988; 88EP-00870008.
XX 30-JAN-1987; 87US-00009325.
XX (SMIK) SMITH KLINE-RIT.
XX (SKFK) SMITH KLINE-RIT.
XX Cabazon T, De Wilde M, Harford N;
XX WPI; 1988-229751/33.
XX N-PSDB; AAN81108.
XX DNA encoding hepatitis B virus antigens and hybrids contg. them - used
PT for expression in yeast to obtain vaccines and bivalent vaccines.
XX Example; Fig 3Aa-3Af; 101pp; English.
XX Sequence of the CS gene (AAN81108) is from lambda-mpf1. A recombinant DNA
CC molecule is claimed, comprising functional DNA coding sequence fused, in
CC phase, to a portion of the Pre-S2 region of a hepatitis B virus (HBV) Pre
CC -S2-S protein coding sequence. The functional DNA coding sequence
CC comprises the Pre-S2 coding sequence, Pre-S1 coding sequence or entire
CC Pre-S1-Pre-S2 coding sequence, the CS protein coding sequence of
CC Plasmodium, or a HIV coding sequence such as an HIV envelope gene
CC sequence, e.g. HIV C7 protein coding region, HIV Peptide 121 coding
CC region, or HIV Dreesman peptide coding region. (Updated on 25-MAR-2003 to
XX correct PA field.)
SQ Sequence 411 AA;

Query Match 37.4%; Score 99; DB 1; Length 411;
Best Local Similarity 56.1%; Pred. No. 0.00032;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;
QY 7 LRPSSGSPSLD-----EKKIAKMEKASSVFNVNSSG 39
: : : : :
Db 360 IKPGSANKPKDLDYNDIEKKICKMEKCSSVFNVNSSIG 400

Search completed: March 10, 2004, 09:12:14
Job time : 76.7977 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 09:16:59 ; Search time 40.0856 Seconds
(without alignments)
268.645 Million cell updates/sec

Title: US-09-848-834A-20

Perfect score: 285

Sequence: 1 XHWYGLRPGSSGSLDEKK.....NVNSSSGPSLHWSYGLRXPX 51

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	263	99.2	51	9	US-09-848-834A-20
2	182	68.7	36	9	US-09-848-834A-16
3	160.5	60.6	46	9	US-09-848-834A-19
4	157	59.2	33	9	US-09-848-834A-12
5	144	54.3	47	9	US-09-848-834A-17
6	138.5	52.3	50	9	US-09-848-834A-18
7	99	37.4	396	10	US-09-820-843A-31
8	97.5	36.8	49	9	US-09-019-010-4
9	97.5	36.8	49	10	US-09-305-924-11
10	97.5	36.8	695	10	US-09-305-924-13
11	95	35.8	20	9	US-09-848-834A-3
12	89	33.6	19	14	US-10-239-313A-54
13	89	33.6	21	10	US-09-932-165-1482
14	89	33.6	21	10	US-09-935-384-710
15	89	33.6	21	10	US-09-942-052-711

Sequence 1404, Ap
Sequence 97, Appl
Sequence 33, Appl
Sequence 761, App
Sequence 761, App
Sequence 652, App
Sequence 26, Appl
Sequence 20, Appl
Sequence 2583, Ap
Sequence 25, Appl
Sequence 2, Appl
Sequence 52, Appl
Sequence 2, Appl
Sequence 44, Appl
Sequence 4224, Ap
Sequence 39, Appl
Sequence 15, Appl
Sequence 35, Appl
Sequence 34, Appl
Sequence 14, Appl
Sequence 13, Appl
Sequence 26, Appl
Sequence 29, Appl
Sequence 30, Appl
Sequence 31, Appl
Sequence 23, Appl
Sequence 15, Appl
Sequence 17, Appl
Sequence 39, Appl

US-10-001-469-1404
US-10-128-711-97
US-10-116-118-33
US-10-082-109A-761
US-10-005-480A-761
US-10-277-292-652
US-10-291-241-26
US-10-280-340-652
US-10-099-460-20
US-10-024-652-2583
US-10-120-885A-25
US-10-107-532-2
US-10-121-016-52
US-10-114-669-2
US-10-120-835-44
US-10-149-138-4224
US-10-114-432-39
US-09-848-834A-15
US-09-848-201A-35
US-09-964-201A-34
US-09-848-834A-13
US-09-848-834A-14
US-09-964-201A-26
US-09-964-201A-29
US-09-964-201A-30
US-09-964-201A-31
US-09-747-802-23
US-09-865-294-15
US-10-411-544-17
US-10-411-544-39

ALIGNMENTS

RESULT 1

US-09-848-834A-20
; Sequence 20, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent version 3.0
; SEQ ID NO 20
; TYPE: PRT
; LENGTH: 51
; FEATURE:
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 378-398 of Plasmid
; OTHER INFORMATION: ium falciparum circumsporozoite (CSP) protein
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD RES
; LOCATION: (51)..(51)
; OTHER INFORMATION: Amidated glycine or glycineamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(36)
; OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium falciparum
; OTHER INFORMATION: circumsporozoite (CSP) protein

```

; NAME/KEY: PEPTIDE
; LOCATION: (37)..(42)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (43)..(51)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-20

Query Match          99.2%; Score 263; DB 9; Length 51;
Best Local Similarity 100.0%; Pred. No. 2.1e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2 HWSYGLRPGSGPSLDKXIAKMEKASSVFNVNSSGSPSLHWSYGLRP 50
    |||||
Db  2 HWSYGLRPGSGPSLDKXIAKMEKASSVFNVNSSGSPSLHWSYGLRP 50

RESULT 2
US-09-848-834A-16
; Sequence 16, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 16
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Gn
; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 378-398 of
; OTHER INFORMATION: the Plasmodium falciparum circumsporozoite (CSP) protein
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(31)
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor
; OTHER INFORMATION: (Tentoxylisin)
; NAME/KEY: PEPTIDE
; LOCATION: (32)..(37)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (38)..(46)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-19
; Sequence 19, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 19
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of Tetanu
; OTHER INFORMATION: toxoid precursor (Tentoxylisin) linked by a spacer to amino acid
; OTHER INFORMATION: sequence 1-10 of GnRH
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Amidated glycine or glycynamide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(46)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(31)
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor
; OTHER INFORMATION: (Tentoxylisin)
; NAME/KEY: PEPTIDE
; LOCATION: (32)..(37)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (38)..(46)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-12
; Sequence 12, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 12
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 378-398 of th
; OTHER INFORMATION: Plasmodium falciparum CSP protein linked by a spacer to amino ac
; OTHER INFORMATION: id sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Amidated aspartic acid
; NAME/KEY: MOD_RES

```

```

Query Match          68.7%; Score 182; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.3e-16;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2 HWSYGLRPGSGPSLDKXIAKMEKASSVFNVNSSGSPSLHWSYGLRP 36
    |||||
Db  2 HWSYGLRPGSGPSLDKXIAKMEKASSVFNVNSSGSPSLHWSYGLRP 36

RESULT 3
US-09-848-834A-19
; Sequence 19, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A

```

```

; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 19
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of Tetanu
; OTHER INFORMATION: toxoid precursor (Tentoxylisin) linked by a spacer to amino acid
; OTHER INFORMATION: sequence 1-10 of GnRH
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Amidated glycine or glycynamide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(46)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(31)
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor
; OTHER INFORMATION: (Tentoxylisin)
; NAME/KEY: PEPTIDE
; LOCATION: (32)..(37)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (38)..(46)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-19
; Sequence 19, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 19
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 378-398 of th
; OTHER INFORMATION: Plasmodium falciparum CSP protein linked by a spacer to amino ac
; OTHER INFORMATION: id sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Amidated aspartic acid
; NAME/KEY: MOD_RES

```

```

Query Match          60.6%; Score 160.5; DB 9; Length 46;
Best Local Similarity 66.0%; Pred. No. 2.5e-13;
Matches 33; Conservative 3; Mismatches 7; Indels 7; Gaps 2;

Qy  2 HWSYGLRPGSGPSLDKXIAKMEKASSVFNVNSSGSPSLHWSYGLRP 50
    |||||
Db  2 HWSYGLRPGSGPSLDKXIAKMEKASSVFNVNSSGSPSLHWSYGLRP 45

RESULT 4
US-09-848-834A-12
; Sequence 12, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 12
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 378-398 of th
; OTHER INFORMATION: Plasmodium falciparum CSP protein linked by a spacer to amino ac
; OTHER INFORMATION: id sequence 2-10 of the GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Amidated aspartic acid
; NAME/KEY: MOD_RES

```

```
; LOCATION: (33)..(33)
; OTHER INFORMATION: Amidated glycine or glycylamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(20)
; OTHER INFORMATION: Amino acid sequence 378-398 of the Malaria
; OTHER INFORMATION: (Plasmodium falciparum) circumsporozoite
; NAME/KEY: PEPTIDE
; LOCATION: (21)..(24)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (25)..(33)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-12

Query Match          59.2%; Score 157; DB 9; Length 33;
Best Local Similarity 94.1%; Pred. No. 4.8e-13;
Matches 32; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 17 DEKIAKMEKASSVFNVNSSGSPSLHWSYGLRP 50
      |||||
DB 1 DEKIAKMEKASSVFNVN--SGPSLHWSYGLRP 32

RESULT 5
US-09-848-834A-17
; Sequence 17, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aptton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 17
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Gn
; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of
; OTHER INFORMATION: the Measles virus protein F linked by a spacer to amino acid seq
; OTHER INFORMATION: uence 2-10 of the GnRH hormone
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD RES
; LOCATION: (47)..(47)
; OTHER INFORMATION: Amidated-glycine or glycylamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(18)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (19)..(34)
; OTHER INFORMATION: Amino acid sequence 288-302 of the Measles virus fusion protein,
; NAME/KEY: PEPTIDE
; LOCATION: (35)..(38)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (39)..(47)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-17

Query Match          54.3%; Score 144; DB 9; Length 47;
Best Local Similarity 57.1%; Pred. No. 3.4e-11;
Matches 28; Conservative 8; Mismatches 9; Indels 4; Gaps 2;
```

```
QY 2 HWSYGLRPGSSGSPSLDEKIAKMEKASSVFNVNSSGSPSLHWSYGLRP 50
      |||||
DB 2 HWSYGLRPGSSGSPSL--KLLSEIK--GVIVHLEGVGSPSLHWSYGLRP 46
```

RESULT 6

```
US-09-848-834A-18
; Sequence 18, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aptton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 18
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
```

```
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of the Te
; OTHER INFORMATION: anus toxoid precursor (Tentoxylisin) protein linked by a spacer
; OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD RES
; LOCATION: (50)..(50)
; OTHER INFORMATION: Amidated glycine or glycylamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
```

```
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(37)
; OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (Ten
; OTHER INFORMATION: oxylisin
; NAME/KEY: PEPTIDE
; LOCATION: (38)..(41)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (42)..(50)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-18
```

```
Query Match          52.3%; Score 138.5; DB 9; Length 50;
Best Local Similarity 54.7%; Pred. No. 1.9e-10;
Matches 29; Conservative 4; Mismatches 11; Indels 9; Gaps 2;
```

```
QY 2 HWSYGLRPGSSGSPSLDEKIAKMEKASSVFNVNSSGSPSLHWSYGLRP 50
      |||||
DB 2 HWSYGLRPGSSGSPSLFNNFTVFWLRVPKVA-----SHLEGPSLHWSYGLRP 49
```

RESULT 7

```
US-09-820-843A-31
; Sequence 31, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTE
; TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
```

; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 31
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Circumsporozoite (CS) protein
; NAME/KEY: misc feature
; OTHER INFORMATION: gi|4493889
US-09-820-843A-31

Query Match 37.4%; Score 99; DB 10; Length 396;
Best Local Similarity 56.1%; Pred. No. 0.00027;
Matches 23; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 7 LRGSSGSPSLD-----EKTKAKMEKASSVFNVNSSSG 39
:|||||
Db 345 IKPGSANKPKDELVDYANDIEKKICKMEKSSVFNVNSSIG 385
:|||||

RESULT 8
US-09-019-010-4
; Sequence 4, Application US/09019010
; Patent No. US20010014330A1
; GENERAL INFORMATION:
; APPLICANT: HARLAND, RICHARD
; APPLICANT: MANN, JOHN G.
; APPLICANT: ACRES, STEPHEN D.
; TITLE OF INVENTION: IMMUNIZATION AGAINST ENDOGENOUS
; MOLECULES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBINS & ASSOCIATES
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/019,010
; FILING DATE: 05-FEB-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,883
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0035
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 325-7812
; TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-019-010-4

Query Match 36.8%; Score 97.5; DB 9; Length 49;
Best Local Similarity 42.9%; Pred. No. 3.3e-05;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

QY 2 HWSYGLRPGSSGFSLDEKTKAKMEKASSVFNVNSSSGPSLHWSYGLRP 50
:|||||
Db 2 HWSYGLRPGSSGFSLHWSYGLRP 50
:|||||

Db 2 HWSYGLRPGS-----GSQDWSYGLRPGSSQHSYGLRP 35

RESULT 9
US-09-305-924-11
; Sequence 11, Application US/09305924A
; Publication No. US20030091579A1
; GENERAL INFORMATION:
; APPLICANT: Jack G. Manns
; APPLICANT: Stephen D. Acres
; APPLICANT: Richard Harland
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
; FILE REFERENCE: 9001-0048
; CURRENT APPLICATION NUMBER: US/09/305,924A
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,217
; EARLIER FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 49
; TYPE: PRT
; ORGANISM: GnRH
US-09-305-924-11

Query Match 36.8%; Score 97.5; DB 10; Length 49;
Best Local Similarity 42.9%; Pred. No. 3.3e-05;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

QY 2 HWSYGLRPGSSGFSLDEKTKAKMEKASSVFNVNSSSGPSLHWSYGLRP 50
:|||||
Db 2 HWSYGLRPGS-----GSQDWSYGLRPGSSQHSYGLRP 35
:|||||

RESULT 10
US-09-305-924-13
; Sequence 13, Application US/09305924A
; Publication No. US20030091579A1
; GENERAL INFORMATION:
; APPLICANT: Jack G. Manns
; APPLICANT: Stephen D. Acres
; APPLICANT: Richard Harland
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
; FILE REFERENCE: 9001-0048
; CURRENT APPLICATION NUMBER: US/09/305,924A
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,217
; EARLIER FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 695
; TYPE: PRT
; ORGANISM: GnRH
US-09-305-924-13

Query Match 36.8%; Score 97.5; DB 10; Length 695;
Best Local Similarity 42.9%; Pred. No. 0.00082;
Matches 21; Conservative 2; Mismatches 11; Indels 15; Gaps 1;

QY 2 HWSYGLRPGSSGFSLDEKTKAKMEKASSVFNVNSSSGPSLHWSYGLRP 50
:|||||
Db 10 HWSYGLRPGS-----GSQDWSYGLRPGSSQHSYGLRP 43
:|||||

RESULT 11
US-09-848-834A-3
; Sequence 3, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047